KU Leuven
Biomedical Sciences Group

Faculty of Kinesiology and Rehabilitation Sciences
Department of Neuromotor Rehabilitation
Pediatric Research Unit

CONSENSUS-BASED CLASSIFICATION OF GAIT IN CHILDREN WITH CEREBRAL PALSY

Angela NIEUWENHUYS

Leuven, September 2016

Dissertation presented in partial fulfillment of the requirements for the degree of Doctor in Biomedical Sciences

Jury:

Promoter:	Prof. Dr. Kaat Desloovere
Co-promoter:	Prof. Dr. Ir. Tinne De Laet
	Prof. Dr. Guy Molenaers
Chair:	Prof. Dr. Hilde Feys
Jury members:	Prof. Dr. Ilse Jonkers
	Prof. Dr. Nathalie Goemans
	Prof. Dr. Ir. Herman Bruyninckx
	Prof. Sylvia Õunpuu

Cover: Lisa Goelen

Table of Contents

List of frequently used abbreviations

3DGA	Three-dimensional gait analysis
BTX-A	Botulinum Toxin type A
CI	Confidence interval
CP	Cerebral palsy
GMFCS	Gross motor function classification scale
POA	Percentage of agreement
ROM	Range of motion
SD	Standard deviation
SPM	Statistical parametric mapping
TD	Typically developing
PS	Pelvis in sagittal plane
HS	Hip in sagittal plane
KSTS	Knee during stance in sagittal plane
KSWS	Knee during swing in sagittal plane
ASTS	Ankle during stance in sagittal plane
ASWS	Ankle during swing in sagittal plane
PC	Pelvis in coronal plane
HC	Hip in coronal plane
PT	Pelvis in transverse plane
HT	Hip in transverse plane
FPA	Foot progression angle

Acknowledgements

Beste lezer, bijna vier jaar geleden nam ik voor het eerst plaats aan het bureau in UZ Pellenberg waar ik nu de laatste hand leg aan het proefschrift dat voor u ligt. Nog maar net mijn Masterdiploma op zak en als echte 'rookie', begon ik toen de krijtlijnen voor vier jaar onderzoek uit te tekenen. Ik hoop dat het resultaat aan uw verwachtingen zal voldoen, maar vooral hoop ik, dat u zich tijdens het lezen bewust bent van de steun die vele collega's en vrienden mij onvoorwaardelijk gegeven hebben. Hun invloed is voelbaar in het design van elke studie, in de opbouw van elke paragraaf, alsook in de professionele en persoonlijke competenties die ik de afgelopen jaren verder heb mogen ontwikkelen.

In het bijzonder erken ik de hulp van mijn promotoren alsook de toewijding die zij aan de dag leggen bij de begeleiding van hun doctorandi: Prof. Kaat Desloovere (Faculteit Bewegings- en Revalidatiewetenschappen (FaBeR)), prof. Tinne De Laet (Faculteit Ingenieurswetenschappen) en prof. Guy Molenaers (Faculteit Geneeskunde).
Elke vergadering bood me nieuwe inzichten, hielp me om prioriteiten te stellen, of leidde me naar één van de vele contactpersonen uit hun netwerk, die me zouden helpen de uitdagingen op te lossen die opdoken tijdens de verschillende studies.

In dit proefschrift vindt u twee internationale studies terug, waaraan in totaal meer dan 40 chirurgen, kinesitherapeuten, kinderartsen, en bewegingswetenschappers deelgenomen hebben. Zij spendeerden ondanks hun volle agenda's verschillende uren aan dit project en ik dank hen voor hun enthousiaste en vrijwillige inzet. Zonder hen geen data, geen resultaten, geen proefschrift.

Een grote 'dankjewel' richt ik ook aan Leen Van Gestel. Ongeveer acht jaar geleden stond zij aan de wieg van deze onderzoekslijn omtrent gangclassificatie bij kinderen met cerebrale parese. Haar input in mijn eerste onderzoeksjaar bij de opbouw en uitvoering van de consensus studie (hoofdstuk 3), is cruciaal geweest om het project te doen slagen. Ook Enrico Di Lello en Erwin Aertbeliën (Faculteit Ingenieurswetenschappen) ben ik dankbaar om de eerste stappen te zetten naar de automatisatie van de ontwikkelde classificatie.

In het begin van mijn tweede onderzoeksjaar leerde ik 'one-dimensional statistical parametric mapping' kennen of kortweg 'SPM', een statistische methode die in hoofdstuk 2 en hoofdstuk 5 van dit werk werd toegepast. Het heeft meer dan een jaar geduurd vooraleer ik me comfortabel voelde om de analyses zelf uit te voeren en te interpreteren. Ik ben veel dank

verschuldigd aan prof. Todd Pataky (Shinshu University, Japan), die Skype-sessie na Skype-sessie rustig uitlegde hoe SPM werkt, en voor welke onderzoeksvragen de methode geschikt is. In één adem wil ik ook collega's Kevin Deschamps en Christophe Meyer bedanken voor de vele brainstorm-sessies, alsook het vertrouwen dat zij me gegeven hebben om als clinicus aan de slag te gaan met complexe statistische methoden en een studie met een belangrijke methodologische component uit te voeren.

As a co-promoter, I was blessed to mentor Eirini Papageorgiou, who was the first physical therapy student choosing this research line on gait classification to pursue her Master's thesis. Out of all thesis students I have guided, she was the one continuously pushing me to reflect on my choices. With every *"But why do you do this?"* she has unconsciously, yet significantly improved the quality of the design and execution of the reliability study (chapter 4). I can think of no better friend and colleague to continue our research and I wish her all the best in completing her PhD a few years from now.

In addition, I would like to say a heartfelt 'thank you' to all colleagues at the gait lab in Pellenberg. With numerous little things, you have, each in your own way, contributed to the end result of this thesis. In particular, I am grateful to Marije Goudriaan, for every little Matlab script and for always 'having my back' in the office - both literally and figuratively.

Daarnaast erken ik graag alle feedback die de examencommissie mij gegeven heeft. Dankzij de kritische vragen van prof. Ilse Jonkers (FaBeR), prof. Herman Bruyninckx (Faculteit Ingenieurswetenschappen), prof. Sylvia Õunpuu (University of Connecticut, USA), prof. Nathalie Goemans (Faculteit Geneeskunde) en prof. Hilde Feys (FaBeR), werd ik gestimuleerd mijn project in een breder onderzoeksdomein te plaatsen en mijn belangrijkste bevindingen en conclusies helder te formuleren.

Voor lof betreffende de cover van dit proefschrift, verwijs ik u met veel plezier door naar Lisa Goelen. Daarenboven zijn taalfouten uit den boze! Veel dank aan Maarten Van Dyck voor de talrijke uren die gespendeerd werden aan het editen van alle teksten.

Het is moeilijk uit te drukken in woorden welke impact Elia Luyten en Luc Nieuwenhuys, in het kort 'mama en papa', hebben gehad op dit proefschrift. Ik draag dit werk dan ook aan hen op, omdat zij er altijd voor mij zijn, omdat geen twee mensen het meer verdienen dan zij.

Angela Nieuwenhuys 23 September 2016, Leuven

Summary

The most common physical disability in children is cerebral palsy (CP). The prevalence of this neuromotor disorder is estimated at 1.7 per 1000 live births. CP is caused by non-progressive brain lesions that occur during the antenatal, perinatal, or postnatal period, at a time when the brain and spinal cord are not yet fully developed. As brain lesions are not curable, treatment mainly focusses on symptom management. Primary motor symptoms in CP are spasticity, weakness, impaired balance, and loss of selective motor control. These symptoms often have a destructive influence on a patient's ability to walk, which is one of the most crucial functional activities of daily life. In CP, about 70% of children are able to walk, albeit with major or minor gait deviations and with or without the use of walking aids. The clinical presentation of gait in CP is ever changing due to the complex interplay of those primary motor symptoms with a maturing brain, growth, and treatment. As a result, secondary symptoms such as muscle contractures and lever arm dysfunctions will eventually occur, for which invasive treatments such as orthopedic surgery are required. To guide treatment planning, gait is typically evaluated through instrumented, three-dimensional motion analysis (3DGA), which provides a highly detailed assessment of joint angles, joint moments, and power during walking. The challenge with using this comprehensive biomechanical measurement of gait is the clinical interpretation of the vast amount of multidimensional data that it generates. The benefits of 3DGA as opposed to observational gait analysis are therefore dependent on the expertise of the clinical professional who is analyzing the data. To this date, there is no standardized method to qualitatively interpret 3DGA data, and there is a lack of effective and robust tools that capture the full complexity of gait reliably and validly, with widespread clinical acceptance and applicability. In routine clinical and research practice, the amount of 3DGA data is reduced before it is analyzed and interpreted. Two approaches for reduction are commonly applied. The first approach analyses gait features, which are specific points extracted from the kinematic and kinetic waveforms. Within the scope of this PhD research, important steps were undertaken in the search for an alternative, standardized method to extract and analyze the clinically relevant information from 3DGA data (**study 1**). The second approach to reduce 3DGA data is to define gait patterns, which allocate multiple gait features, either within a joint or across multiple joints, into groups. The principal goal of the PhD was to develop a clinically relevant, reliable, and valid classification for pathological movement patterns during gait in children with CP (**study 2, 3, 4, 5**).

In the **first study**, a literature review found that approximately 220 papers have reported on the outcome of treatment in CP by evaluating children pre- and post-treatment using 3DGA features. Focusing on the studies that evaluated the effect of Botulinum Toxin type A (a common treatment intervention to manage spasticity), this first study shows that there is no consensus regarding the selection of gait features that are expected to be sensitive to change after treatment in CP. Feature analysis may fail to provide a full understanding of the effect of treatment. Clinically relevant information could be missed, as the selection of features is most likely based on the clinical expertise of the medical team that decides on the treatment plan. In a subsequent retrospective intervention study, statistical parametric mapping (SPM) was identified as a valid, unbiased statistical alternative. SPM allowed kinematic and kinetic waveforms to be evaluated as a whole. This statistical approach eliminates the need for a priori data reduction, and keeps the probability of making a Type I or Type II error stable by considering the interdependency of all points of the waveform.

Regarding gait classifications in CP, a literature review showed that several gait patterns and classifications based on kinematic and kinetic data have been previously reported in literature. However, their clinical applicability is limited because psychometric properties of reliability and validity are often not yet established. In the **second study**, a Delphi consensus project was organized to ensure that the new classification was clinically relevant. The Delphi approach is a semi-quantitative research method where an international expert panel was consulted via iterative surveys to provide their opinion on the problem of gait classification in CP. The study started with a first proposal of gait patterns that should be included in the classification, based on previous literature and on the expertise of the clinical and research team of University Hospitals Leuven. After three consecutive survey rounds, consensus was reached on 49 gait patterns across the pelvis, hip, knee, and ankle joints in the sagittal, coronal, and transverse plane.

After the development of the classification, a necessary next step was to ensure that the level of clinician agreement on the patterns of a patient was at a sufficiently high level so that the patterns could be used reliably in practice. In the **third study**, an international agreement study was conducted among 29 clinicians with varying levels of experience with regards to CP and 3DGA. Apart from a few individual patterns, the study demonstrated that, after a brief learning phase, clinicians could assign the consensus-based joint patterns with good consistency between and within raters. The amount of patient data that were found to be

‘unclassifiable’ by the clinicians was low compared to previously published classifications and formed an indirect confirmation of the content validity of the patterns.

As the patterns defined during the Delphi study were the result of an informed, yet subjective opinion of an expert panel, the classification might have still provided an incomplete picture on gait pathology in CP. Therefore, the **fourth study** examined the content validity of the classification. It was assessed whether objective patient data supported the existence of the patterns. To this end, SPM was used to analyze a large database of classified kinematic and kinetic trials. This was to confirm whether each of the 49 patterns differed from the gait pattern of typically developing children in the key areas of the gait cycle that were indicated in the pattern definitions by the experts. Even though this hypothesis was largely confirmed, some additional areas that were not included within the definitions of the Delphi patterns were highlighted by the SPM analysis.

The **fifth study** examined the construct validity of the classification, measuring the extent to which the gait patterns of the classification system are able to distinguish between the categories of other validated scales that measure the same or a related construct. Therefore, the prevalence of the patterns in a large cohort of children with CP was evaluated. It was found that the distribution of these patterns was associated with the distribution of other relevant and validated scales in CP, such as topographical classification, gross motor function, and levels of spasticity and weakness.

In conclusion, this PhD research has made important contributions to the analysis and standardized interpretation of 3DGA data. The classification that was developed within this research presents clinicians and researchers with a comprehensive overview of clinically relevant gait features and gait patterns, which will hopefully serve as a basis for improved communication and a more uniform terminology regarding gait pathology in CP. Patterns and their definitions should be adapted when necessary, and future research should further demonstrate the validity, responsiveness, and clinical applicability of the classification. One of the main contributions of this PhD research is that the developed methodological framework, combining qualitative and quantitative research methods to build a clinically relevant, reliable, and valid classification, could also be generalized to any other medical condition that affects movement.

Samenvatting

Hersenverlamming of cerebrale parese (CP) is de vaakst voorkomende lichamelijke handicap bij kinderen. De prevalentie van deze neuromotorische aandoening wordt geschat op 1.7 per 1000 levendgeborenen. De oorzaak van CP zijn niet-progressieve hersenletsels die zich voordoen tijdens de prenatale, perinatale of postnatale periode, op het moment dat de hersenen en het ruggenmerg nog niet volledig ontwikkeld zijn. Omdat deze hersenletsels ongeneesbaar zijn, focust de behandeling van CP zich hoofdzakelijk op de symptomen. Primaire motorische symptomen in CP zijn spasticiteit, spierzwakte, verstoord evenwicht en verminderde selectieve motorische controle. Deze symptomen hebben vaak een destructieve invloed op het vermogen van de patiënt om te stappen, wat één van de meest cruciale functionele activiteiten in het dagelijks leven is. Ongeveer 70% van de kinderen met CP kan stappen, zij het met grote of kleine pathologische gangafwijkingen, en met of zonder het gebruik van loophulpmiddelen. Het gangpatroon van kinderen met CP is voortdurend in verandering, door de complexe interactie van de primaire motorische symptonen en de zich ontwikkelende hersenen, groei en behandeling. Hierdoor ontstaan secundaire symptomen zoals spiercontracturen en benige deformiteiten, waardoor invasieve behandelingen zoals orthopedische chirurgie noodzakelijk zijn. Een evaluatie aan de hand van geïnstrumenteerde, drie-dimensionele ganganalyse (3DGA) wordt doorgaans gebruikt ter ondersteuning van het plannen van behandelingen. Tijdens 3DGA wordt een zeer gedetailleerde meting gemaakt van gewrichtshoeken, alsook van de momentwerking en het vermogen rond de verschillende gewrichten tijdens het stappen. De grote uitdaging bij het gebruik van deze uitgebreide biomechanische gangmeting, is de klinische interpretatie van de enorme hoeveelheid multidimensionale data die gegenereerd wordt. Daardoor zijn de voordelen van 3DGA, ten opzichte van ganganalyse door (video)observatie, afhankelijk van de expertise van de klinische expert die de data analyseert. Op dit moment is er geen gestandaardiseerde methode om 3DGA data op een kwalitatieve manier te interpreteren. Daarnaast is er een tekort aan effectieve en robuuste methoden die de volledige complexiteit van gang op een betrouwbare en valide manier incorporeren en bovendien kunnen bogen op wijdverspreide klinische erkenning en toepasbaarheid. Zowel in de klinische praktijk als voor de onderzoekswereld, is het gebruikelijk de hoeveelheid 3DGA data te reduceren alvorens ze te analyseren en interpreteren. Deze datareductie wordt gewoonlijk op twee manieren bekomen. De eerste methode definieert gangkarakteristieken of ‘features’. Dit zijn specifieke punten die uit de

kinematische en kinetische curves worden geëxtraheerd. Dit doctoraatsonderzoek zet belangrijke stappen in de zoektocht naar een alternatieve, gestandaardiseerde methode om de klinisch belangrlijke informatie uit 3DGA te extraheren en analyseren (studie 1). Een tweede mogelijke methode om 3DGA data te reduceren omvat het definiëren van gangpatronen. Dergelijke gangpatronen groeperen meerdere gangkarakteristieken, ofwel binnen één gewricht, ofwel over meerdere gewrichten heen. Het hoofddoel van het doctoraat bestond erin een klinisch relevante, betrouwbare, en valide classificatie te ontwikkelen voor pathologische bewegingspatronen tijdens het stappen bij kinderen met CP (**studie 2, 3, 4, 5**).

In de **eerste studie** heeft een literatuuroverzicht aangetoond dat ongeveer 220 publicaties reeds rapporteerden over de effectiviteit van een behandeling bij kinderen met CP door de analyse van gangkarakteristieken voor en na de behandeling. Na gedetailleerde evaluatie van de publicaties die het effect van Botulinum Toxine type A behandeling rapporteerden (d.i. een veelvuldig gebruikt middel om spasticiteit te behandelen), werd besloten dat er geen consensus bestaat met betrekking tot de gangkarakteristieken waarvan men verwacht dat ze een verbetering in het gangpatroon zullen aantonen na behandeling. De analyse aan de hand van gangkarakteristieken zal waarschijnlijk niet volstaan om een volledig begrip te krijgen van de veranderingen in het gangpatroon na behandeling. Klinisch relevante informatie kan gemist worden omdat de selectie van de features afhangt van de subjectieve, klinische expertise van het medische team dat beslist over het behandelingsplan. In een hierop volgende retrospectieve interventiestudie werd ‘statistical parametric mapping’ (SPM) geïdentificeerd als een valide, statistisch alternatief. Deze benadering is meer vrij van bias omdat het mogelijk is de kinematische en kinetsche curves in één geheel te analyseren. Deze statistische methode maakt a priori datareductie aan de hand van gangkarakteristieken dus overbodig, en zorgt ervoor dat de waarschijnlijkheid om een Type I of Type II fout te maken stabiel blijft, door de onderlinge tijdsafhankelijkheid van alle punten van een curve in aanmerking te nemen.

In wetenschappelijke literatuur werd reeds een waaier van gangclassificaties gepubliceerd, die veelal gebaseerd zijn op kinematische en kinetische data. Hun klinische toepasbaarheid is echter beperkt gebleven, aangezien hun betrouwbaarheid en validiteit vaak nog niet bewezen is. In de **tweede studie** werd een Delphi-consensus-project georganiseerd, om erover te waken dat de nieuwe classificatie klinisch relevant zou zijn en om een goede inhoudelijke validiteit na te streven. De Delphi aanpak is een semi-kwantitatieve onderzoeksmethode waarbij een internationaal panel van experten wordt geconsulteerd via herhaaldelijke

vragenlijsten, om aan te geven welke de klinisch relevante gangpatronen zijn bij kinderen met CP. De studie startte met een voorstel van gangpatronen en definities die in de classificatie zouden moeten worden opgenomen, gebaseerd op de bestaande literatuur en de kennis van experten die deel uitmaken van de onderzoeksgroep en het klinische team van de Universitaire Ziekenhuizen te Leuven. Na drie opeenvolgende vragenlijsten (Delphi rondes) werd een consensus bereikt over 49 gangpatronen van de pelvis-, heup-, knie- en enkelgewrichten in het sagittale, coronale en transversale vlak.

Nadat de classificatie ontwikkeld was, bestond de volgende noodzakelijke stap eruit om na te gaan of clinici die de classificatie gebruiken consistent dezelfde gangpatronen definiëren bij dezelfde patiënten, opdat men de classificatie met vertrouwen zou kunnen toepassen in de praktijk. In de **derde studie** werd daarom een internationale betrouwbaarheidsstudie uitgevoerd tussen 29 clinici met wisselende ervaring wat betreft CP en 3DGA. Na een korte leerfase werd voor de gehele classificatie een goede betrouwbaarheid gemeten, met uitzondering van een aantal specifieke patronen. De hoeveelheid patiënten die door de clinici beoordeeld werden als “onclassificeerbaar”, was laag in vergelijking met eerder gepubliceerde studies. Dit vormde een indirect bewijs voor de inhoudelijke validiteit van de patronen.

Omdat de patronen die gedefinieerd werden tijdens het Delphi-consensus-project het resultaat waren van een geïnformeerde, doch subjectieve mening van een panel van experten, is het mogelijk dat de patronen een onvolledig beeld van de gangpatronen in CP weergeven. Daarom onderzocht de **vierde studie** de inhoudelijke validiteit van het classificatiesysteem. Er werd bestudeerd of het bestaan van de 49 gedefinieerde patronen aangetoond kan worden door een objectieve, statistische analyse van 3DGA data van patiënten met CP. Hiertoe werd door middel van SPM een grote database geanalyseerd die geclassificeerde kinematische en kinetische grafieken bevat. Deze analyse evalueerde in welke mate elk van de 49 opgestelde patronen afweek van het gangpatroon van normaal ontwikkelende kinderen in die kerngebieden van de gangcyclus die gedefinieerd waren door de experten tijdens de consensus studie. Hoewel deze hypothese grotendeels bevestigd kon worden, identificeerde de SPM analyse ook bijkomende gebieden die niet in de definities van de Delphi patronen voorkwamen.

De **vijfde studie** onderzocht de constructvaliditeit van het classificatiesysteem, die nagaat in welke mate de patronen van het classificatiesystem onderscheid kunnen maken tussen de categorieën van andere gevalideerde schalen die een gerelateerd construct meten. Hiertoe werd de prevalentie van de patronen in een groot cohort van kinderen met CP geëvalueerd.

Deze studie toonde aan dat de distributie van de gangpatronen geassocieerd was met de distributie van andere relevante en gevalideerde schalen in CP, zoals topografische classificatie, grove functionele motoriek, en niveaus van spasticiteit en spierzwakte.

Dit doctoraatsonderzoek heeft waardevolle bijdragen geleverd aan de analyse en gestandaardiseerde interpretatie van 3DGA data. De ontwikkelde classificatie biedt clinici en onderzoekers een uitgebreid overzicht van de klinisch relevante gangkarakteristieken en gangpatronen bij kinderen met CP. De patronen zullen een basis vormen voor een verbeterde communicatie en het gebruik van een meer uniforme terminologie betreffende de kenmerken van pathologische gang bij CP. Patronen en hun definities moeten aangepast worden indien nodig en toekomstig onderzoek zal de validiteit, de responsiviteit, en de klinische toepasbaarheid van de classificatie verder moeten bevestigen. Een fundamentele bijdrage van dit doctoraatsonderzoek is dat het ontwikkelde methodologische kader, waarbij gebruik gemaakt werd van het gezamelijk potentieel van gedegen kwalitatieve en kwantitatieve onderzoeksmethoden om een klinisch relevante, betrouwbare en valide classificatie op de bouwen, toegepast kan worden bij elke medische aandoening die het menselijk bewegen verstoort.

Chapter 1

Introduction

Background

This general introductory chapter offers the reader the necessary background information for the different research studies composing this PhD thesis. The thesis focusses on the analysis and interpretation of gait measurements in children with cerebral palsy (CP). The first part of the introduction briefly discusses the definition and epidemiology of CP, as well as its relevant clinical subtypes and classifications. Subsequently, three-dimensional gait analysis (3DGA) is introduced as a golden standard to evaluate pathological gait in CP. The state of the art concerning gait features and existing gait classification systems, which are based on 3DGA data, are further presented. The final part states the research aims, demonstrates the cohesion of the chapters within the thesis, and describes the retrospective database, which contains the experimental population for the different research studies of the thesis.

This doctoral thesis relates to a larger research path which originated in 2007 at the Department of Rehabilitation Sciences at KU Leuven and the Clinical Motion Analysis Laboratory at University Hospitals Leuven, in Belgium[1]. Since 2012, this project is funded by KU Leuven (OT/12/100) and constitutes a shared collaboration between researchers of the Department of Rehabilitation Sciences and the Department of Mechanical Engineering. **The overall aim of the project is to develop a clinically relevant and automated classification system for pathological gait in children with CP, by improving the state of the art on the analysis and classification of continuous waveforms from a clinical and engineering perspective**. From the clinical perspective, the presently introduced doctoral thesis primarily focused on the development of a clinical gait classification in CP and the exploration of its reliability and validity. From the engineering perspective, probabilistic methods to automate clinical classifications using continuous waveform data were explored[2].

Cerebral palsy

Definition

There has been much debate about the definition of CP. In mid-nineteenth century, William John Little, Jakob von Heine, and Eduard Heinrich Henoch were the first to describe a group of disorders in children, which is nowadays recognized as CP, but was then named 'Little's disease'[3–6]. The term 'cerebral palsy' was later introduced in 1889 by William Osler, who reported 151 case series in his work: '*The Cerebral Palsies of Childhood'*[7]. During the following decades, CP remained ill-defined and constituted a large umbrella term for a variety of childhood disorders. An important evolution in the search for a universally accepted definition took place around 1950, when the American Academy for Cerebral Palsy suggested to restrict the definition of CP to non-progressive neurological disorders and thereby excluding all progressive neuromotor conditions and brain neoplasms[4,8]. A new definition by Bax in 1964[9] was updated during an international workshop in 2004, and resulted in the most comprehensive definition to date, which defines CP as *'a group of disorders of the development of movement and posture, causing activity limitation, that are attributed to non-progressive disturbances that occurred in the developing fetal or infant brain. The motor disorders of cerebral palsy are often accompanied by disturbances of sensation, cognition, communication, perception, and/or behavior, and/or by a seizure disorder.'*[10]. Hence, CP remains a clinical, rather than an etiological diagnosis. Although not without debate[11], the abovementioned definition of CP is widely applied throughout medical and research centers.

Epidemiology

Prevalence

CP is the most common physical disability in children. Its prevalence in registries around the world ranges between 1.5 to 3.5 per 1000 live births[12–15]. Over the past decades, these numbers have remained generally stable, although fluctuations have been noted. At the end of the 1970s and beginning of 1980s, an increase in prevalence of CP was noted as a result of advancements in neonatal care that ensured a higher survival rate of premature infants[16]. Since then, a slight but significant decrease has been observed in the prevalence of CP in low birth weight infants in European and Australian registries[17–19]. The prevalence of CP has also been reported to be up to 30% higher in boys compared to girls[14]. In part, this could be due to

an increased susceptibility in boys for white matter injury and aberrant fetus growth during pregnancy[20,21]. Exact mechanisms underlying these gender differences remain to be unraveled.

Pathological brain patterns

Each element of the abovementioned definition of CP is indicative of the wide range of different clinical disorders in CP, yet the emphasis primarily lies with the impairment of movement and posture. To better understand the variety of movement impairments and more specifically gait pathology in CP, it is briefly reviewed how movement is regulated in typically developing children and healthy adults. To walk, or to perform any kind of movement, the human body requires above all the ability to contract muscles. Muscle contractions are initiated and coordinated by the neuromotor system, which is typically divided in an upper and lower motor neuron system. The upper motor neuron system consists of four neural components or area's in the brain that are interconnected and that systematically plan, regulate, and adapt movement[22]:

1) The **cortical motor centers** consist of the supplementary motor cortex, the premotor cortex, and primary motor cortex. In these centers, the thought of motion arises and the commands to move are issued via the corticospinal tract to the lower motor neurons in the spinal cord.
2) The **basal ganglia** consist of five nuclei which contain 'memories' of previous movement with respect to the location of the body in space. This information is relayed to the cortical motor centers before the command to move is given. The basal ganglia therefore have an important role in the planning and initiation of movements.
3) Next, there is the **cerebellum**, which plays a crucial role in balance. The cerebellum also monitors if a movement is executed according to the commands given by the cortical motor centers. For this task, the cerebellum relies on the sensory information coming from the muscles via the spinocerebellar tract. Based on this information, proposals to adapt and correct ongoing movements are passed on to the cortical motor centers.
4) Two important **nuclei in the brainstem** are the vestibular nucleus and the reticular formation. These nuclei provide the cortical motor centers with proprioceptive information on the posture and tonus of the trunk and of the proximal limbs.

The lower motor neuron system consists of motor neurons that are located in the anterior horn of the spinal cord. These motor neurons transfer commands from the upper motor neuron system via peripheral nerves to the muscle and back.

CP is an upper motor neuron syndrome and develops when lesions occur in any of those key areas before the neural system has fully developed. Primary motor symptoms that occur as a result of brain lesions are spasticity, muscle weakness, disturbed balance, and loss of selective motor control[23]. Based on the nature, location, extent, and timing of the brain lesion, the clinical presentation of CP will vary. This also implies that CP has many causal pathways. Even though neuroimaging techniques, and more specifically magnetic resonance imaging (MRI), are not considered essential in the diagnosis of CP, they have been crucial in making progress toward a better understanding of the mechanisms of brain injury[16,24–26]. Authors have reported that more than 80% of children with CP present with pathological neuroimaging findings[25,27].

A reliable, standardized classification of pathological brain lesions based on MRI imaging was developed to be used by clinicians and researchers. This classification identifies brain lesions according to their time of occurrence during brain development (Figure 1)[24]. There are two main phases in brain development; the first phase is the phase of cortical neurogenesis, during which precursor neuronal cells proliferate and migrate to organize the cerebral cortex. Disturbances during this period cause brain malformations, which have been observed in approximately 10% of children with CP[26]. The second phase of neural development starts with the final trimester of pregnancy. This phase continues after birth and is characterized by axon and dendrite growth, and myelination. Destructive processes causing brain injury during this period are typically predominantly white or grey matter injuries. Periventricular white matter injury is found in more than 50% of children with CP and is usually discovered in patients with spastic cerebral palsy, affecting both sides of the body[16,24]. Combinations of grey and white matter injuries also occur[24,26]. In approximately 10-20% of children with CP, MRI findings are normal[26].

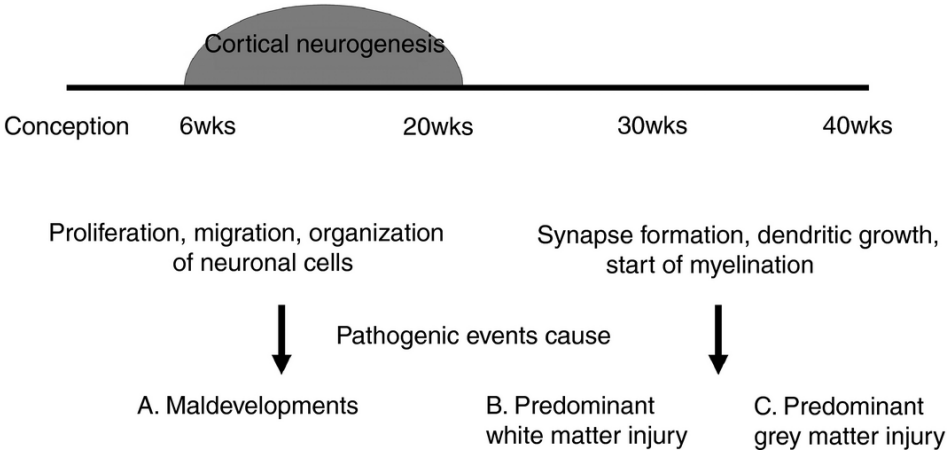


Figure 1. Overview of pathogenic patterns in CP based on timing of occurrence during brain development. Reprinted by permission from Mac Keith Press. *Dev Med Child Neurol*, copyright 2016[24].

Classification of clinical subtypes in CP

Even though brain lesions are non-progressive, the clinical appearance of CP is always changing because it is continuously altered by a maturing brain, musculoskeletal growth, and various treatment therapies[10]. Because of this wide heterogeneous presentation, it has long been of interest to classify the clinical characteristics of CP[4,8]. In 1998, the 'Surveillance of Cerebral Palsy in Europe' (SCPE) was founded, constituting a multi-center European collaboration between more than twenty centers across eight countries[28]. The SCPE has made important contributions toward the standardized definition and classification of CP and its clinical subtypes[29]. A classification was proposed, based on the clinical findings in multiple domains, among which there were for instance epilepsy, and visual and hearing impairments. The categorizations relevant to posture and abnormal movement patterns are related to neurological motor type, topographical distribution of symptoms, and gross and fine functional motor ability.

Neurological motor type

Based on neurological signs, three types of CP were defined: spastic, dyskinetic, and ataxic type of CP[29]. Spasticity refers to "a velocity dependent increase in hypertonia with a catch when a threshold is exceeded", and is by far the most prevalent motor type in CP, with over 80% of all children being classified into this group[14,30]. The other types of CP are far less common and constitute approximately 15% of the CP population[14,18]. Dyskinetic CP is further

divided in a dystonic and choreo-athetotic group, which are characterized by uncontrolled, jerky movements. Ataxic CP is least common and is known for impaired coordination, loss of balance control, tremor, and hypotonia. Gait in the dyskinetic and ataxic CP groups is characteristically inconsistent. If patients present with a combination of these neurological features, it is suggested that the dominant feature determines the group in which a patient is classified.

Topographical classification

A topographical classification is also often referred to. This system classifies patients based on the parts of the body that are affected in patients with spastic CP (Figure 2). For years, a distinction was made between hemiplegia, diplegia, and quadriplegia, but also other forms such as monoplegia and triplegia were not uncommon. These classifications have caused many concerns and are not as straightforward to apply as they may seem. The SCPE therefore suggested a simplified categorization of 'unilateral' vs. 'bilateral' spastic CP[29]. To date, this categorization remains controversial, however its reliability has been found to be good to excellent and as such, it was also adopted for the research studies conducted within this PhD thesis[31–33]. Approximately 35% of all children with spastic CP are in the group that is unilaterally affected[14,34].

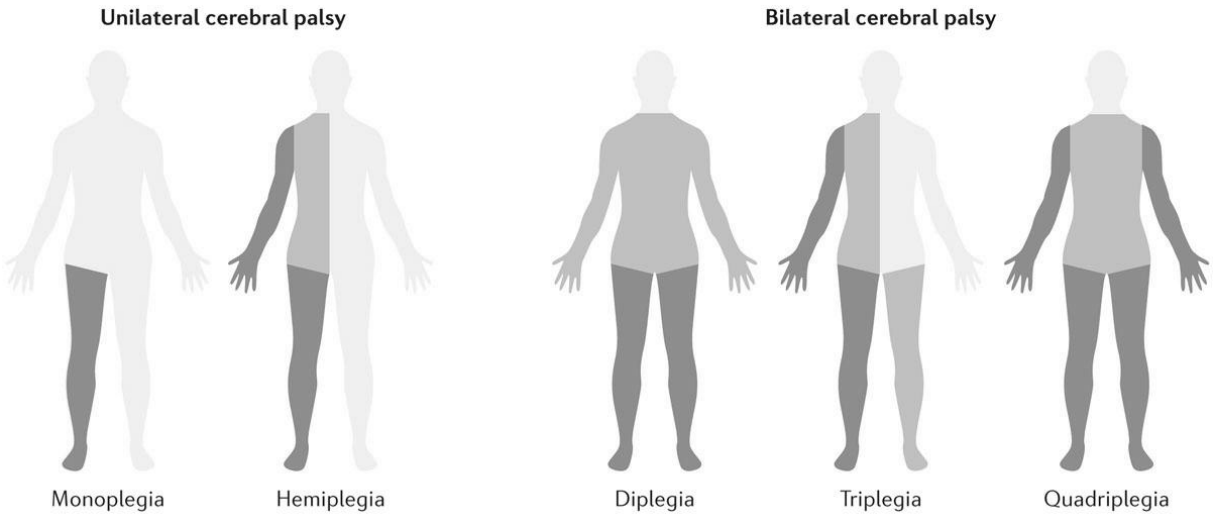


Figure 2. Topographical classifications in CP. Reprinted by permission from Macmillan Publishers Ltd. *Nat Rev Dis Primers*, copyright 2016[16].

Gross and fine functional motor ability

Several classifications of functional motor ability have been developed for use in CP. Two scales are often used to describe gross motor function, namely the Functional Mobility Scale (FMS) and the Gross Motor Function Classification System (GMFCS)[35,36]. The GMFCS is used most often and describes the ability of a child to sit, to perform transfers, or to walk, run, and jump[36]. It is a five-point ordinal scale with separate descriptions for different age groups. In general, children with GMFCS level I can perform most activities such as walking and climbing stairs independently with minimal restrictions. GMFCS level III is characterized by the need for walking aids such as crutches or walkers, whereas children with GMFCS level V are typically dependent on manual or powered wheelchairs for mobility, and require physical assistance for all activities. Morris and Bartlett[37] have presented a literature review, showing how the GMFCS has been adopted in observational and experimental research, for instance to explore how GMFCS relates other relevant impairments and subgroups in CP, or to select and describe study samples. The reliability of the GMFCS has been documented extensively, as well as assessments of its responsiveness, and its content, construct, and cross-cultural validity. The authors concluded that the GMFCS has established itself as the principal classification system of functional ability for children with CP, demonstrating sound validity and good reliability for children between 2-12 years old[37]. Similar to the GMFCS, the Manual Ability Classification System (MACS) and Bimanual Fine Motor Function (BFMF) are reported to describe upper limb function (fine motor abilities) in CP[38,39].

Gait

In CP, about 70% of children are able to walk, albeit with major or minor pathological deviations from normal, and with or without the use of walking aids[34]. The heterogenic clinical presentation of CP is especially striking when looking at gait. Therefore, many attempts to classify this wide variety of gait deviations have been reported before[40]. The following will discuss different purposes of classification and introduce the normal gait pattern of healthy, unimpaired individuals. Subsequently, the golden standard to evaluate gait, 3DGA, will be introduced, before focusing on the analysis of features and gait classifications based on 3DGA data that have been previously reported in literature.

Purposes and scope of gait classification

Classification has been defined as *"the act or process of putting people or things into groups based on ways that they are alike"*, or as *"a systematic arrangement in groups or categories according to established criteria"* (www.merriam-webster.com). One of the main purposes for creating classifications is that several medical diagnoses, as is the case with CP, do not provide sufficient information to make treatment plans, nor can they provide adequate information on expected functional outcome[41,42]. Developing valid and reliable tools to measure, analyze, and interpret pathological gait in CP are therefore essential steps to successfully develop and implement patient-specific treatment plans. Classifications of gait in CP can support medical practitioners in their clinical reasoning, as clinicians can compare the gait pattern of a new patient with the patterns of previous patients to help decide which treatment might be optimal. As was described by Bax et al.[10], classifications typically have four purposes: (a) description, (b) prediction, (c) comparison, and (d) evaluation of change. Translated to the problem of gait pathology in CP, this means a good classification should allow clinicians to (a) describe the quality of gait and its relevant deviations, (b) make prognoses on how a subjects' gait will evolve over time and support treatment planning, (c) compare gait between patients or between groups of patients, and (d) permit evaluation of changes in gait over time or after treatment. Such a classification would facilitate communication, not only between healthcare workers, but also between healthcare workers and their patients or patients' families[40,43]. A uniform terminology with regards to relevant gait patterns in children with CP would also have other advantages. Firstly, classifications

would provide an objective tool for research purposes, and if used to describe study populations, would allow for a better and more transparent interpretation of scientific literature[44]. Hence, it would ease the comparison of outcomes across different studies. Secondly, classifications will likely simplify sharing and merging of patient data across multiple research centers and thereby facilitate multicenter studies[28]. The potential of classifications to have such effects has been shown by the wide uptake of the GMFCS in both clinical practice and academia[37]. Lastly, classifications would be useful to health care professionals who are learning to interpret gait biomechanics in CP.

The newly developed classification within this PhD thesis intends to be relevant for all ambulatory children with spastic CP. To be able to serve the purposes described above, the classification should be aimed at defining patterns that allow a detailed interpretation of the quality of gait. Hence, it is not meant to be a functional classification, such as the GMFCS, which focuses on severity of functional limitations. Functional classifications differ from the intended classification of the PhD thesis in that they cannot cover qualitative information on the movement of each individual joint. They are therefore less suited to facilitate clinical reasoning. Considering the International Classification of Functioning, Disability and Health[41] model, the GMFCS is located at the level of activity restrictions, while the classification put forward by in this PhD thesis is situated closer to the level of body impairment and structure.

Typical gait

The normal gait pattern is a continuous and consistent series of gait cycles. A gait cycles starts and ends with the heel contact of one foot with the ground (Figure 3)[45]. Typically, a gait cycle is divided in a stance phase, (i.e. 0-60% of the gait cycle) and a swing phase (i.e. 60-100% of the gait cycle). The stance phase begins and ends with a period of double support, during which both feet touch the ground. The first double support period is called ‘loading response’, while the second is known as ‘pre swing’ phase. During gait, loading response of the right leg coincides with the pre-swing phase of the left and vice versa. Single stance is the period between loading response and pre swing, and this phase is further split up into ‘midstance’ and ‘terminal stance’. These phases combined coincide with the swing phase of the opposite limb. Swing phase is divided in initial swing, mid swing, and terminal swing. Apart from indicating specific events to define spatio-temporal parameters such as walking velocity or

step length, these phases are traditionally used to describe the motion of the different lower limb joints throughout the gait cycle.

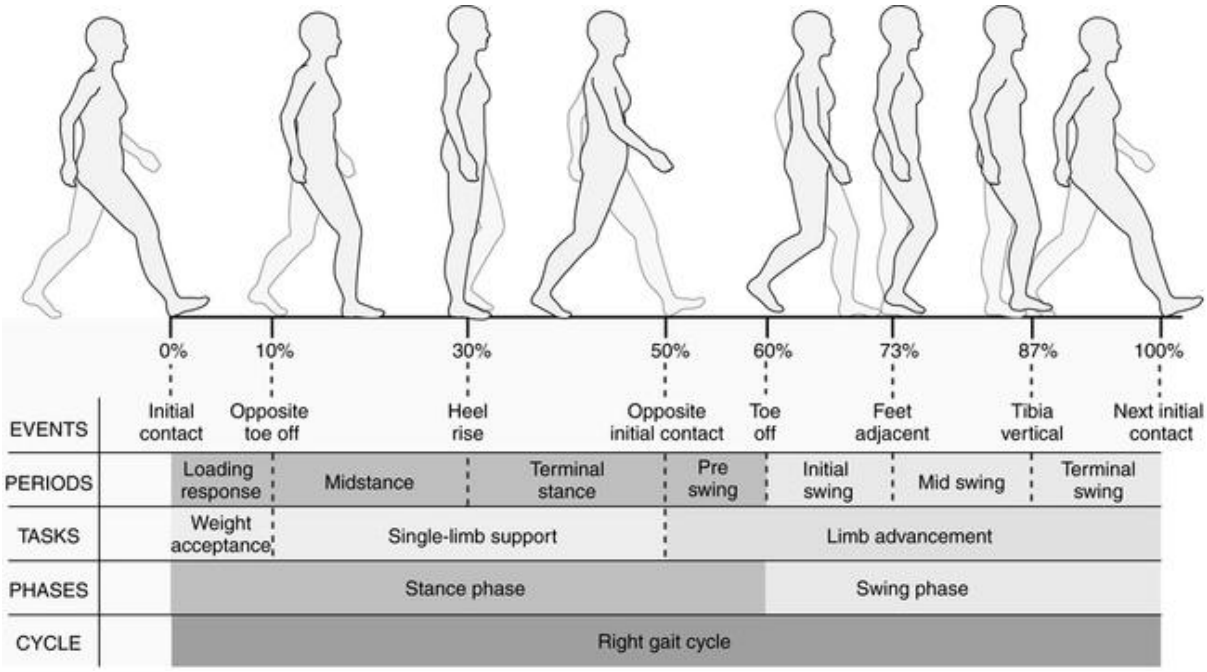


Figure 3. Overview of the different phases of the gait cycle. (Figure adopted from Neumann DA, Kinesiology of the musculoskeletal system: foundations for physical rehabilitation, St. Louis, Mosby, 2010).

Evaluation of gait

To evaluate gait in children with CP, there are two frequently applied approaches: observational gait analysis (i.e. examining gait during clinical examination or via video recordings) and instrumented 3DGA. Observational gait analysis has the advantage of being easily applicable, as well as being very time- and cost-effective. Because observational gait analysis is subjective in nature, several tools such as the Edinburgh Visual Gait Score[46] have been developed to standardize the interpretation procedure[47–50]. However, none of these tools achieve the objectivity, reliability, or validity of 3DGA[51,52]. 3DGA provides a highly detailed assessment of the motion of the different lower limb joints during the abovementioned phases of gait in the sagittal, coronal, and transverse plane. To this end, optoelectronic cameras measure the movements of retro-reflective markers that are placed on anatomical landmarks on the skin, which allow calculating kinematic data (i.e. joint angles and muscle lengths). Force platforms measure the ground reaction force and allow the calculation of kinetic data

(i.e. joint moments and power). Kinematic and kinetic data are typically segmented and time-normalized to gait cycles. Kinetic data is commonly normalized to body mass.

Because of its objectivity and high level of accuracy, 3DGA has become the golden standard to evaluate gait. For instance, with the help of 3DGA, surgery in CP has evolved towards a stronger focus on the correction of bony deformities and away from lengthening of soft tissue contractures such as the Achilles tendon and hamstrings lengthening[53–55]. Several authors have shown that the clinical decision making is significantly altered by incorporating 3DGA in the clinical decision making process and that patient outcomes improve more when treatments follow the recommendations based on 3DGA[56–60]. It is apparent that the definition of 'improved patient outcome' based on 3DGA data is different across these studies. This also illustrates the main challenge of using this comprehensive biomechanical measurement of gait: how does one analyze and interpret the kinematic and kinetic data in a standardized, clinically meaningful way? The clinical interpretation and analysis of the vast amount of multidimensional 3DGA is puzzling and subjective[61,62]. On top of the high dimensionality of the data, 3DGA waveforms are time-dependent and correlated with each other, and relations between different gait waveforms are non-linear[62]. In addition to these factors, clinicians should also be aware of potential measurement errors that could influence outcome and should thus be taken into account when interpreting the results[63]. In routine clinical and research practice, the analysis and interpretation of 3DGA data is facilitated by data reduction. Generally, data reduction is accomplished via (a) the calculation of summary indices, (b) the analysis of gait features, or (c) the definition of gait patterns (i.e. gait classification). Indices, such as the Gait Deviations Index[64], Gait Profile Score[65], or Movement Deviation Profile[66] are measures of severity of gait pathology, which quantify the distance between a patient's gait from normal. Unlike gait features and gait classifications, these summary measures do not allow an interpretation of the quality of gait (i.e. which joints are deviating in which direction?). Therefore, indices are less suited to assist the clinical decision making process.

State of the art on gait features and gait classifications

Gait features

The analysis of gait features is more frequently discussed in literature than gait patterns or classifications. Within the context of this PhD thesis, gait features refer to a specific point of a kinematic or kinetic waveform, e.g. peak values. The usefulness of 3DGA features for clinical and research practice, is dependent on their potential to contribute to clinical decision making, which is determined by their responsiveness to change over time or after treatment. In research, gait features are commonly used to analyze the effect of a specific treatment approach on gait or to identify underlying clinical causes of pathological gait deviations[67,68]. There are three problems related to the analysis of gait features. Firstly, it is often unclear in literature how, why, and who decides on the features that should be taken into consideration in statistical analysis. It is assumed that feature selection is a subjective process, which is therefore dependent on the clinical expertise of the involved clinicians or researchers. Due to this subjective selection without the report of a clearly directed hypothesis, results could be biased and potentially relevant discriminatory features risk being omitted[69]. A solution for this issue was proposed by Wolf et al.[70], who presented a methodological framework for automatic detection of the most relevant clinical features, dependent on the clinical research question. Secondly, feature definitions are typically based on the behavior of normal kinematic and kinetic waveforms. If the shape of the pathological kinematic data is heavily different from TD gait, some features are less meaningful or difficult to extract[62]. Thirdly, statistical approaches to evaluate these features usually involve the analysis of multiple dependent gait features extracted from each waveform. It is challenging to account for the dependency of these features during statistical analysis without needlessly decreasing power (i.e. the ability to detect a significant effect if it is present in the data)[71,72]. Alternatively, by failing to account for the dependency between features, the risks of false positive outcomes can drastically increase, especially with small study sample sizes[73,74]. Given that feature analysis suffers from these disadvantages, it is necessary to search for alternative approaches that are better suited to answer biomechanical hypotheses, which typically pertain to full kinematic or kinetic waveforms (or phases of these waveforms) as opposed to specific features. Statistical parametric mapping, or SPM, has emerged as a promising statistical alternative for the analysis of multidimensional biomechanical waveform data[69,75]. SPM

performs hypothesis testing in a continuous manner, thereby avoiding the need for subjective, a priori data reduction.

Gait classifications

Attempts to define gait patterns in CP and create gait classifications using 3DGA data have been reported since the 1970s. Simon et al.[76] were first to identify three groups of children displaying knee hyperextension during stance phase. In 2007, Dobson et al.[40] reported on the methodological quality of eighteen gait classifications in a systematic literature review and concluded that *"although gait classification in CP can be useful in clinical and research settings, the methodological limitations of many classifications restrict their clinical and research applicability"*. Only four of those eighteen studies were not based, at least in part, on 3DGA data. As Dobson's review included classifications that were published until March 2005, their systematic search strategy for paper identification was adopted during the course of this PhD research, to systematically review gait classifications that were published from April 2005 until February 4, 2016. The same inclusion and exclusion criteria of Dobson et al.[40] were used, except only full papers were included as the relevant methodological aspects are difficult to fully describe in an abstract. In addition to papers that reported on new gait classifications, separate studies examining the reliability or validity of a previously published gait classification were also included. Twenty-seven studies were identified (Figure 4), among which fourteen studies were considered as 'new' classifications and thirteen studies were building on previously published classifications and/or evaluating their reliability or validity[77–103]. All but three of those fourteen new classifications, were based, at least in part, on kinematic or kinetic data[77,78,92].

The following paragraphs discuss different methodological aspects relevant to gait classifications in CP, namely generalizability, classification development process, reliability, content validity, criterion and construct validity, and responsiveness. The concepts of generalizability, reliability, validity, and responsiveness are interpreted based on the definition of measurement properties as reported in the COSMIN manual and checklist (**CO**nsensus-based **S**tandards for the selection of health **M**easurement **IN**struments)[104,105].

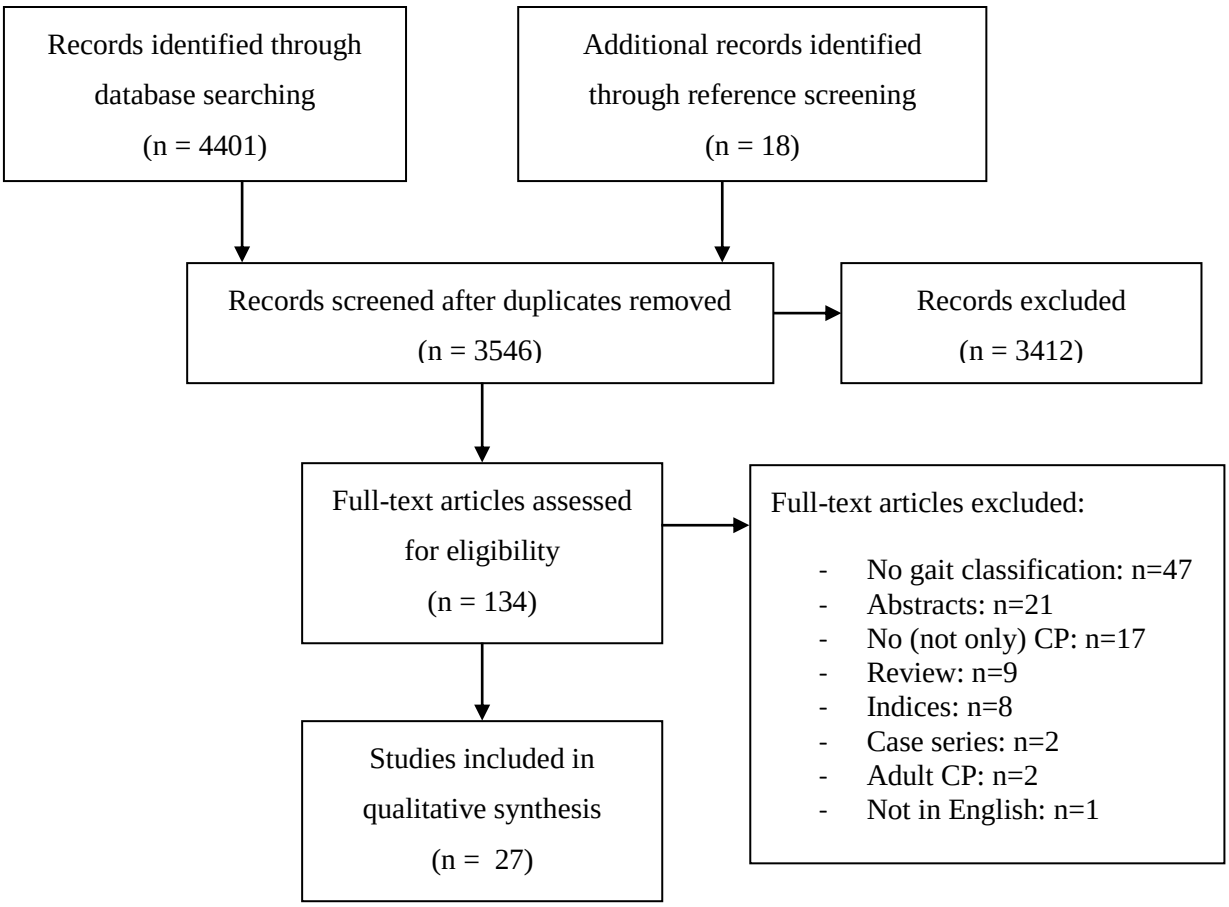


Figure 4. Systematic literature search on the development and psychometric properties of gait classification systems, published after March 2005. Flowchart adapted from PRISMA guidelines[106].

Patient population – generalizability

It should be clear from the report of a study to what extent the classification can be generalized to the entire CP population. Important aspects of generalizability are a description of the study sample (e.g. motor type, previous treatment, age, GMFCS level) as well as the method and setting of patient recruitment in which the study was conducted[104]. Classifications constructed using a database from a hospital setting might differ from population-based databases. Dobson et al.[40] reported that only five out of eighteen studies had adequately defined the studied patient population. For the fourteen new classifications since 2005, patient populations were mostly well described, except for two studies[92,94]. If a classification is constructed without the use of patient data, such as in Ferrari et al.[101], and Davids and Bagley[85], the characteristics of the population for which the classification is intended, should be clearly reported. It is noted that several newer classifications were intended for both

patients with bilateral or unilateral CP[78,79,82,83,93,107], whereas previously, the majority of classifications (14/18 studies) were created for patients with either unilateral or bilateral CP[40]. All classifications so far have focused on children with the spastic motor type, probably due to the inconsistent gait of children with dystonia and ataxia.

Classification development process

Methods to develop a classification can be qualitative or quantitative. **Quantitative development approaches** use data analysis techniques to identify patterns from 3DGA data. A few examples of these techniques are cluster analysis[79,92,95], linear discriminant analysis[98], support vector machines[108], and Bayesian networks[82,109]. Quantitative techniques are objective and the development process of these classifications is typically very well described. However, the clinical interpretation of the groups that are defined based on these methods is difficult and often, no clear clinical description is provided for each of the identified groups[40]. This limits their applicability to assist in the process of clinical decision making. An important issue for quantitative construction approaches is the sample size of the patient population that is used to create the classification. In total, four studies that were reported after 2005 used 30 patients or less to develop a classification[78,93,95,100]. These small sample sizes are not likely to be representative of the large variability of gait deviations that are inherent to patients with spastic CP, thereby limiting the validity and generalizability of these classifications.

A **qualitative development approach** relies on the judgment of clinical experts to describe and define different gait patterns based on deviations from TD gait using quantitative 3DGA data. Qualitatively developed classifications have the advantage of being more readily applicable by clinicians. Moreover, they have the potential to be automated as definitions rely on quantitative 3DGA data. This was illustrated by Padilla et al.[98], who developed an algorithm to automatically classify the patterns of Winters et al.[110] based on sagittal plane knee kinematics. On the other hand, qualitative development approaches are criticized because of their subjectivity and obscurities in reporting the process of classification development, which therefore limits their reproducibility[40]. The best known examples of qualitative classifications are the classification of Winters et al.[110] and of Sutherland et al.[111]. Typically, these classifications represent the vision of a few experts related to one gait laboratory or clinical or academic institution. The search for a more structured and documented qualitative approach, representing consensus among multiple centers is pressing,

and might facilitate a more widespread uptake of a classification in the clinical and research field.

Reliability

For a classification to be useful in clinical practice, it should allow clinicians to consistently recognize the patterns with a sufficient level of agreement. Reliability was only examined by two authors in the review of Dobson et al.[40]. Recently more studies have reported the level of clinician agreement on the classifications of Winters et al.[110] and Rodda et al.[112], using either 3DGA data, video data, or a combination of both[87,90,91]. Out of the fourteen new gait classifications, only Ferrari et al.[81] have provided inter-observer reliability levels for their classification of spastic diplegia and found good to excellent inter-rater agreement, which also included raters who were not experienced with the classification. Generally, intra- and interrater agreement results vary from fair to excellent. Nevertheless, the reliability of the vast majority of classifications remains to be examined.

Content validity

Content validity can be defined as *"the degree to which the content of a classification is an adequate reflection of the construct to be measured"*[104]. It is difficult to give a precise guideline on what constitutes 'good content validity' for gait classifications in CP. The scope of CP gait classifications ranges from 'single joint patterns' to 'multiple joint patterns'. Single joint patterns can sometimes be based on one specific gait feature, but are usually a combination of different features at the level of one joint, as is the case for instance with the knee patterns of Sutherland et al.[111]. Multiple joint patterns combine deviations across more than one joint and/or more than one plane, such as for instance the patterns of Ferrari et al.[101]. If the construct of a classification is "gait", it is recommended that all lower limb joints (and even the trunk) across the three anatomical planes are taken into account, as children with CP are known to show pathological deviations at all levels in all planes.

The content validity of many classifications is jeopardized as only one of eighteen studies identified by Dobson et al.[40] defined gait patterns that included deviations in the sagittal, coronal, and transverse plane. For the novel classifications since 2005, eleven (out of fourteen) were based on 3DGA data and five of them discussed all three anatomical

planes[83,85,95,100,101]. Another threat to content validity was reported by several authors, who discovered 'unclassifiable' patients when examining the classifications of Winters et al.[110] or Rodda et al.[112] [87,88,90,97]. Lastly, several studies using a quantitative classification approach performed a cross-validation of their outcome by examining whether a set of test patients, not included in the classification development process, can be classified into the defined gait patterns[40,82,100]. These analyses can also be considered an indication of content validity. However, the outcome of cross-validation analyses was not always satisfactory[113] and methodological shortcomings hindered their validity[40].

Criterion validity, construct validity

Criterion validity is *"the degree to which an instrument is an adequate reflection of a gold standard"*[104]. Zwick et al.[114] were the only authors to report on the criterion validity of their classification, which differentiated between dynamic tightness and fixed contractures in patients with equinus gait based on subjective evaluation of ankle kinematic and kinetic data. The golden standard involved a clinical examination of passive ankle dorsiflexion under anaesthesia. However, for most classifications such a golden standard is not available as gait patterns are commonly not directly linked to an isolated underlying clinical symptom as was the case in Zwick et al.[114]. By lack of a golden standard, it is appropriate to report on the construct validity of a classification. Construct validity is defined as *"the degree to which scores of an instrument are consistent with hypotheses, for instance with regard to relationships to scores of other instruments or differences between relevant groups"*[104]. Rozumalski et al.[94] investigated how different crouch patterns, which were determined via k-means cluster analysis, were characterized by range of motion, muscle strength, and spasticity. Dobson et al.[86] reported how the distribution of the patterns of Winters et al.[110] was associated with other validated classifications such as the GMFCS and Functional Mobility Scale[35]. Bonnefoy-Mazure et al.[83] and Ferrari et al.[80] have conducted similar research, reporting the extent to which the gait patterns of their classifications were characterized, among others, by spasticity and weakness. Following the COSMIN guidelines, a limitation for these studies is that they fail to state specific, a priori hypotheses, including magnitude and direction of expected associations or correlations with other scales or instruments[105]. For several gait classifications, developing specific, a priori stated hypotheses might be difficult

as it is not yet entirely clear how clinical symptoms such as muscle weakness or spasticity manifest during gait[67,115–117].

Responsiveness

The definition of responsiveness is *"the ability of an instrument to detect change over time in the construct to be measured"*[104]. Responsiveness of gait classifications has been evaluated pre- and post-treatment, for instance after surgery or with the use of orthoses[76,77,88,112,118]. Even though an assessment of responsiveness is important to demonstrate the clinical applicability and relevance of a classification, it is not often examined. Moreover, if responsiveness is evaluated, generalizability of the results is limited because very small sample sizes of between two and fifteen patients were often analyzed[76,77,118]. Riad et al.[88] classified 31 children pre- and post-surgery using the classification of Winters et al.[110], and found that approximately 75% of patients had changed, of which the majority improved towards a pattern with less severe gait deviations. However, a limitation for all these studies is that changes in gait patterns with increasing age or after treatment were never associated or correlated with a golden standard, or with another comparator instrument that evaluates or quantifies changes in gait after treatment. Admittedly, this is a difficult matter, as there is no 'golden standard gait classification' in CP and the definition of success of treatment based on 3DGA data is challenging. In this respect, the Movement Analysis Profile or Gait Profile Score[65] might be an interesting tool, as it provides a quantitative score for the 'normality' of gait at the level of each joint based on 3DGA data and minimal clinically important differences for the tool have recently been defined[119].

Objectives and thesis outline

Objectives

3DGA is considered to be the golden standard to measure gait objectively and with high accuracy. The trigger for this research project was the variability in, and difficulty of interpreting and analyzing the large amount of 3DGA data, such that it is useful in clinical practice[61]. To this date, data reduction methods fail to capture the full complexity of gait pathology in CP reliably and validly, with widespread clinical acceptance and applicability[40,62]. **The principal goal of this PhD research was to develop a clinically relevant, valid, and reliable classification system for pathological movement patterns during gait in children with spastic CP, based on kinematic and kinetic data.**

To achieve this overall goal, five sub-goals were specified as follows:

I. To create a synopsis of the current state of the art with regard to the analysis and interpretation of 3DGA features that are sensitive to treatment, and to explore SPM as a valid and unbiased alternative for statistical analysis of 3DGA data.

II. To develop a new gait classification for children with spastic CP, by achieving an international expert consensus on clinically relevant joint patterns during gait, taking into consideration the available knowledge from literature.

III. To establish the reliability of the classification by evaluating the level of clinician agreement on the joint patterns which were defined during the consensus study.

IV. To explore the content validity of the classification by identifying differences between the consensus-based gait patterns using SPM on classified kinematic and kinetic patient data.

V. To explore the construct validity of the classification by examining the relations between the distribution of the consensus-based gait patterns on one hand, and patient-specific characteristics and clinical symptoms on the other hand.

Outline

Chapters 2 to 6 of this PhD thesis focus on one of the abovementioned sub-goals. Figure 5 outlines the relation between these chapters.

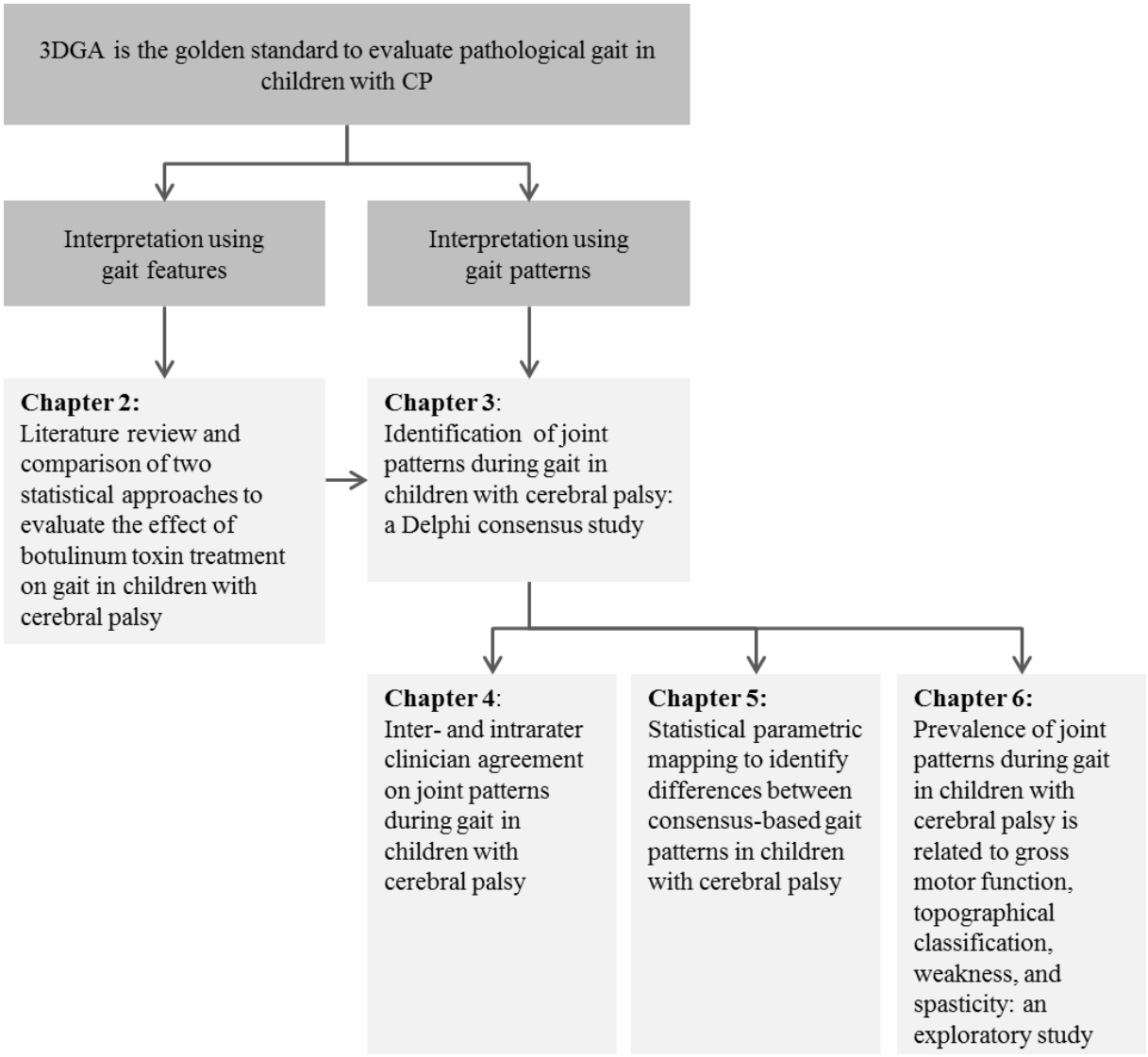


Figure 5. Cohesion between the different research studies within the PhD thesis.

C**hapter 2** was triggered by the different potential sources of bias involved in the traditional kinematic and kinetic feature analysis, which is often applied in CP research to report on the outcome of a treatment intervention. The hypothesis for the study was that studies in literature might over- or underestimate the effect of BTX-A treatment on gait due to (1) statistical approaches that fail to control Type I or Type II error rates and (2) bias within the feature selection process. Chapter 2 describes a literature review, which aimed to provide clinicians

with a reference of clinically relevant kinematic and kinetic features that are sensitive to Botulinum Toxin type A (BTX-A) treatment in children with CP. Subsequently, by comparing two statistical methods to analyze kinematic and kinetic gait data, the outcome of multilevel BTX-A treatment was reported and the value of SPM as an objective and valid alternative statistical approach was examined. To this end, a retrospective sample of 53 children who had undergone 3DGA before and after multilevel BTX-A treatment were recruited from the database of University Hospitals Leuven (Table 2).

Table 2. Demographic characteristics of the CP groups across the different studies and the reference database of typically developing (TD) children.

	TD (N=56)	**Chapter 2 (N=53)**	**Chapter 4 (N=82)**	**Chapter 5 (N=356)**	**Chapter 6 (N=286)**
Gender (n)					
Male	24	18	57	212	165
Female	32	35	25	144	121
Weight (mean (SD), in kg)	40.1 (17.7)	20.1 (7.0)*****	31.5 (15.8)	32.2 (14.0)	34.3 (14.8)
Height (mean (SD), in m)	1.48 (0.21)	1.14 (0.15)*****	1.31 (0.24)	1.34 (0.20)	1.38 (0.20)
Diagnosis (n)					
Bilateral CP	na	36	55	219	166
Unilateral CP	na	17	27	137	120
GMFCS (n)					
Level I	na	25	47	192	172
Level II	na	17	26	117	89
Level III	na	11	9	47	25
Number of 3DGA sessions	56	106	82	459	286
Age at time of 3DGA (mean (SD))	11 ye, 1 mo (3 ye, 10 mo)	6 ye, 1 mo (2 ye, 4 mo)*****	9 ye, 5 mo (3 ye, 11 mo)	9 ye, 10 mo (3 ye, 6 mo)	10 ye, 4 mo (3 ye, 7 mo)

SD = standard deviation; na = not applicable; 3DGA = three-dimensional gait analysis; ye = years, mo = months; ***** = at time of pre-Botulinum Toxin treatment 3DGA.

Chapter 3 describes a Delphi consensus study that was executed to achieve an international expert consensus on the clinically relevant single joint patterns during gait in CP. The hypothesis was that a qualitative, structured consensus approach results in clinically relevant patterns with good face validity. It is expected that consensus among clinicians in the field of gait classification may result in a better uptake of the gait patterns by the research and clinical field, as users may gain more confidence in the patterns, knowing they resulted from a highly experienced, international expert panel. After an initial consensus meeting, iterative online surveys were conducted in search of a consensus on the clinically relevant gait patterns in CP

that can be defined based on the kinematic and kinetic waveforms of the lower limb joints in the sagittal, coronal, and transverse plane.

In a next step, **chapter 4** discusses the reliability of the classification that was developed in the Delphi consensus study by examining the level of clinician agreement with which the patterns could be recognized from 3DGA data. This research study questioned (1) whether clinicians could reliably use the classification and (2) whether clinicians who were experienced with the analysis and interpretation of 3DGA data would achieve higher agreement scores than inexperienced clinicians. The experimental group consisted of 82 patients with CP, recruited from the database of University Hospitals Leuven (Table 2). A clinical rater group was recruited from participant lists of international gait courses in 2015. In the end, inter- and intrarater agreement estimates were calculated based on the ratings of 32 clinicians, who were asked to classify the kinematic and kinetic waveforms of 27 or 28 patients with CP twice, using a custom-made online graphical user interface (www.cmal-tools-leuven.be).

Chapter 5 contributes to the content validity of the classification by evaluating whether the subjective rules that were defined during the consensus study could be confirmed when performing SPM on quantitative, classified kinematic and kinetic patient data. The main hypotheses stated (1) that patterns with no or minor gait deviations do not differ significantly from the gait pattern of TD children, (2) that all other pathological patterns (n=38) differ from TD gait, and (3) that the locations of difference within the gait cycle that are highlighted by SPM, concur with the locations described in the classification rules. Because the research question of this methodological study concerns the analysis of differences between kinematic and kinetic groups as they are defined subjectively by clinicians, all available classified trials (n=1719, n=356 children with CP) were used for statistical analysis (Table 2).

In **chapter 6,** the prevalence and distribution of the consensus-based gait patterns is explored. In addition, the associations between the classification and other existing tools to describe the clinical presentation of CP were investigated. It is hypothesized that the prevalence of the patterns is associated with age, topographical classification, GMFCS level, previous treatment, spasticity, and weakness. This study therefore provided insight toward the construct validity and clinical applicability of the classification (i.e. its potential to describe and compare gait patterns of clinically relevant subgroups of children with CP). For this

study, one gait analysis session was selected for each eligible patient from the retrospective database (n=286), according to specific inclusion- and exclusion criteria (Table 2).

A concluding general discussion (chapter 7) summarizes the results from the different studies and will reflect on methodological decisions and steps for future research.

Database development

As acknowledged before, the work described in this PhD thesis was funded by an OT project of KU Leuven during which the Neuromotor Research group collaborated with the Department of Mechanical Engineering (cfr. supra, 'Background'). Within the engineering department, supervised probabilistic methods to automate clinical classifications were explored[2]. Their methods require a large database containing kinematic and kinetic gait waveforms of children with CP, which are classified by a clinician. At the onset of the project, the development of a large retrospective research database was therefore initiated. The database was set-up to contain demographic patient information (e.g. age, height, and sex), 3DGA data, as well as data from the clinical examination (on muscle contractures, bony deformities, spasticity, muscle weakness, and selective muscle control) that is typically performed along with the 3DGA measurements. Only 3DGA sessions that were conducted by experienced physical therapists, and which were organized in the context of a patient's treatment, were included. A very broad manual search of the database of the clinical motion analysis laboratory of University Hospitals Leuven was conducted to identify gait analysis sessions for patients meeting following inclusion criteria:

(1) A diagnosis of spastic CP,
(2) GMFCS level I, II, or III (i.e. the ability of independent walking), and
(3) Age between 3 and 18 years.

Children with marked signs of dystonia or ataxia were excluded because it is known that gait in these two subgroups is characterized mainly by inconsistency. To confirm the diagnosis of spastic CP retrospectively, the medical records from the hospital were screened. If available, at least five medical reports from three different medical professionals over a period of at least five years were consulted. Children were excluded if marked signs of dystonia or ataxia were present, and if the presence of a medical condition, which could interfere with gait (such as coexisting genetic disorders, severe visual impairment or mental retardation), was identified. The prevalence of marked dystonia and ataxia is approximately 6.5% and 4.3%, which is very

low compared to the overall presence of spasticity in CP[14]. Therefore, the classification developed within this PhD would still be generalizable to the majority of patients with CP.

Because the gait classification is meant to fulfill the purposes of description, prediction, comparison, and evaluation of change in gait in CP[10], the inclusion criteria for the database were defined as broadly as possible. To be able to thoroughly test the clinical applicability of the classification in the future, it was deemed necessary to allow all gait analysis sessions for which good quality kinematic trials could be identified. Eligible gait analysis sessions could therefore be collected for the purposes of re-evaluation (3DGA not specifically related to treatment), pre- or post-surgical evaluations, or pre- or post-BTX-A treatment assessments. All gait analysis sessions fulfilling the criteria between November 2005 and September 2015 were eligible to be included. However, sessions that were used by Van Gestel et al.[1] were also re-examined and included if they fulfilled the criteria and if they were collected after 2000, in order to ensure a stable 3DGA protocol. The main aim of data recruitment was to study as many patients as possible. However, the decision to include multiple gait analysis sessions per patient was also taken to allow future studies to focus on the responsiveness of the classification (i.e. the sensitivity of the patterns to treatment or to change over time). Moreover, for the study described in chapter 2, it was necessary to include more than one session per patient.

Clinical and biomechanical data related to each gait analysis session was anonymised and underwent a data curation process. This ensures that the database is suited for publication, can be further extended, or can be used for future research projects. Sanity tests were performed for all clinical data related to patient demographics, the Modified Ashworth scores for spasticity, and the Manual Muscle Test scores for weakness[120,121]. The quality of each gait trial was also thoroughly inspected. Trials were only included if the patient walked in a straight line and if at least one gait cycle for the left and right leg could be identified. Depending on the date on which the gait session was collected, Vicon Nexus software version 1.8.5, version 1.5.2, or Vicon Workstation version 5.1 (Vicon Motion Systems, Oxford, UK) was used to define gait cycles and to estimate spatio-temporal parameters, joint angles (ankle, knee, hip) or segment orientation (pelvis, foot), internal joint moments in the three anatomical planes, and joint power for the three lower limb joints. Afterwards, custom-made Matlab software was used to screen the quality of each kinematic and kinetic trial. Trials with artifacts or signs of inaccurate marker placement or experimental error were excluded. To this

end, the range of motion and position of the knee varus-valgus angle was evaluated[63]. Outliers, or trials that were not representing a child's gait pattern, were also excluded. Outliers were determined based on visual inspection, when the distance from the trial of a patient to the average of all trials of that patient was larger than two standard deviations (of the TD database).

The database was provided to the Department of Mechanical Engineering for the technical developments of the research project, and was used for the studies described in chapters 2, 4, 5, and 6. Table 2 presents the patient populations for those studies. The full database was used for statistical analysis in chapter 5. Ethical approval for this PhD research was granted by the Medical Ethical Committee of University Hospitals Leuven, reference S56036. Furthermore, the committee approved a research collaboration agreement with prof. Todd Pataky, so that anonymized kinematic and kinetic data could be shared regarding the analyses using SPM described in chapter 2.

References

1. Van Gestel L. Acquired pathological gait patterns and their relationship to structural brain lesions in children with spastic cerebral palsy through Bayesian reasoning: a methodological framework. KU Leuven, 2012.
2. Di Lello E. Bayesian Time-Series Models : Expert Knowledge-Driven Inference and Learning for Engineering Applications. KU Leuven, 2015.
3. Little W. Hospital for the Cure of Deformities: course of lectures on the deformities of the human frame. *Lancet*. 1843;41:350-354.
4. Morris C. Definition and classification of cerebral palsy: a historical perspective. *Dev Med Child Neurol Suppl*. 2007;109:3-7.
5. von Heine J. *Spinale Kinderlabmung*. Stuttgart: Cotta'sche Buch; 1860.
6. Henoch E. "De atrophia cerebri" (translated and summarized). In: *In: Thomson, J. Henoch, E. (1889) Lectures on Children's Disease*. Vol London: Sampson Low; 1842.
7. Osler W. *The Cerebral Palsies of Childhood*. London: HK. Lewis; 1889.
8. Minear W. A classification of cerebral palsy. *Pediatrics*. 1956;18:841-852.
9. Bax M. Terminology and classification of cerebral palsy. *Dev Med Child Neurol*. 1964;6:295-307.
10. Bax M, Goldstein M, Rosenbaum P, et al. Proposed definition and classification of cerebral palsy, April 2005. *Dev Med Child Neurol*. 2005;47(8):571-576.
11. Blair E, Love S. Definition and classification of cerebral palsy Commentary. *Dev Med Child Neurol*. 2005;47:510.
12. Blair E. Epidemiology of the cerebral palsies. *Orthop Clin North Am*. 2010;41(4):441-455.
13. Reid SM, Carlin JB, Reddihough DS. Rates of cerebral palsy in Victoria, Australia, 1970 to

2004: Has there been a change? *Dev Med Child Neurol*. 2011;53(10):907-912.

14. SCPE (surveillance of cerebral palsy in europe). Prevalence and characteristics of children with cerebral palsy in Europe. *Dev Med child Neurol*. 2002;44(9):633-640.
15. Oskoui M, Coutinho F, Dykeman J, Jetté N, Pringsheim T. An update on the prevalence of cerebral palsy: A systematic review and meta-analysis. *Dev Med Child Neurol*. 2013;55(6):509-519.
16. Graham HK, Rosenbaum P, Paneth N, et al. Cerebral palsy. *Nat Rev Dis Prim*. 2016:15082.
17. Sellier E, Platt MJ, Andersen GL, Krägeloh-Mann I, De La Cruz J, Cans C. Decreasing prevalence in cerebral palsy: A multi-site European population-based study, 1980 to 2003. *Dev Med Child Neurol*. 2015:85-92.
18. Australia Cerebral Palsy Research Institute. *Australian Cerebral Palsy Register Report 2013, Birth Years 1993 - 2006*. Sydney; 2013.
19. Platt MJ, Cans C, Johnson A, et al. Trends in cerebral palsy among infants of very low birthweight (<1500 g) or born prematurely (<32 weeks) in 16 European centres: a database study. *Lancet*. 2007;369(9555):43-50.
20. Johnston M V, Hagberg H. Sex and the pathogenesis of cerebral palsy. *Dev Med Child Neurol*. 2007;49(1):74-78.
21. Jarvis S, Glinianaia S V, Arnaud C, et al. Case gender and severity in cerebral palsy varies with intrauterine growth. *Arch Dis Child*. 2005;90(5):474-479.
22. Gage JR, Schwartz MH, Koop SE, Novacheck TF. BThe identification and treatment of gait problems in cerebral palsy. In: Vol 2nd ed. London: Mac Keith Press; 2009:1-19.
23. Gage JR, Schwartz MH, Koop SE, Novacheck TF. CThe identification and treatment of gait problems in cerebral palsy. In: Vol 2nd ed. London: Mac Keith Press; 2009:107-129.
24. Himmelmann K, Horber V, De La Cruz J, et al. MRI classification system (MRICS) for children with cerebral palsy: development, reliability, and recommendations. *Dev Med Child Neurol*. 2016:1-8.
25. Krägeloh-Mann I, Horber V. The role of magnetic resonance imaging in elucidating the pathogenesis of cerebral palsy: A systematic review. *Dev Med Child Neurol*. 2007;49(2):144-151.
26. Korzeniewski SJ, Birbeck G, DeLano MC, Potchen MJ, Paneth N. A systematic review of neuroimaging for cerebral palsy. *J Child Neurol*. 2008;23(2):216-227.
27. Reid SM, Dagia CD, Ditchfield MR, Carlin JB, Reddihough DS. Population-based studies of brain imaging patterns in cerebral palsy. *Dev Med Child Neurol*. 2014;56(3):222-232.
28. Cans C. Surveillance of cerebral palsy in Europe: a collaboration of cerebral palsy surveys and registers. Surveillance of Cerebral Palsy in Europe (SCPE). *Dev Med Child Neurol*. 2000;42(12):816-824.
29. Cans C, Dolk H, Platt M, Colver A, Prasauskiene A, Krägeloh-Mann I. Recommendations from the SCPE collaborative group for defining and classifying cerebral palsy. *Dev Med Child Neurol Suppl*. 2007;109:35-38.
30. Sanger TD, Delgado MR, Gaebler-spira D, Hallett M, Mink JW. Classification and definition of disorders causing hypertonia in childhood. 2003;111(1):e89-e97.
31. Shevell MI. The terms diplegia and quadriplegia should not be abandoned. *Neurology*. 2009:1-2.
32. Colver AF, Sethumadhavan T. The term diplegia should be abandoned. *Ach Dis Child*. 2003;88:286-291.

33. Sellier E, Horber V, Krägeloh-Mann I, De La Cruz J, Cans C. Interrater reliability study of cerebral palsy diagnosis, neurological subtype, and gross motor function. *Dev Med Child Neurol*. 2012;54(9):815-821.

34. Beckung E, Hagberg G, Uldall P, Cans C. Probability of walking in children with cerebral palsy in Europe. *Pediatrics*. 2008;121(1):e187-e192.

35. Graham HK, Harvey A, Rodda J, Nattrass GR, Pirpiris M. The Functional Mobility Scale (FMS). *J Pediatr Orthop*. 2004;24(5):514-520.

36. Palisano R, Rosenbaum P, Walter S, Russell D, Wood E, Galuppi B. Development and reliability of a system to classify gross motor function in children with cerebral palsy. *Dev Med Child Neurol*. 1997;39(2):214-223.

37. Morris C, Bartlett DJ. Gross Motor Function Classification System : impact and utility. *Dev Med Child Neurol*. 2004;46:60-65.

38. Eliasson A-C, Krumlinde-Sundholm L, Rösblad B, et al. The Manual Ability Classification System (MACS) for children with cerebral palsy: scale development and evidence of validity and reliability. *Dev Med Child Neurol*. 2006;48(7):549-554.

39. Beckung E, Hagberg G. Neuroimpairments, activity limitations, and participation restrictions in children with cerebral palsy. *Dev Med Child Neurol*. 2002;44(5):309-316.

40. Dobson F, Morris ME, Baker R, Graham HK. Gait classification in children with cerebral palsy: A systematic review. *Gait Posture*. 2007;25(1):140-152.

41. World Health Organization. Towards a Common Language for Functioning , Disability and Health ICF. *Int Classif*. 2002;1149:1-22.

42. Rosenbaum P, Stewart D. The World Health Organization International Classification of Functioning, Disability, and Health: A Model to Guide Clinical Thinking, Practice and Research in the Field of Cerebral Palsy. *Semin Pediatr Neurol*. 2004;11(1):5-10.

43. Wren TAL, Rethlefsen S, Kay RM. Prevalence of specific gait abnormalities in children with cerebral palsy: influence of cerebral palsy subtype, age, and previous surgery. *J Pediatr Orthop*. 2005;25(1):79-83.

44. Chambers H. The Delphi consensus technique: Oracle of gait analysis. *Dev Med Child Neurol*. 2016;58(3):228.

45. Perry J. *Gait Analysis: Normal and Pathological Function*. Thorofare, NJ, Slack; 1992.

46. Read HS, Hazlewood ME, Hillman SJ, Prescott RJ, Robb JE. Edinburgh Visual Gait Score for Use in Cerebral Palsy. *J Pediatr Orthop*. 2003:296-301.

47. Toro B, Nester CJ, Farren PC. The development and validity of the Salford Gait Tool: An observation-based clinical gait assessment tool. *Arch Phys Med Rehabil*. 2007;88(3):321-327.

48. Kawamura CM, de Morais Filho MC, Barreto MM, de Paula Asa SK, Juliano Y, Novo NF. Comparison between visual and three-dimensional gait analysis in patients with spastic diplegic cerebral palsy. *Gait Posture*. 2007;25(1):18-24.

49. Brown CR, Hillman SJ, Richardson AM, Herman JL, Robb JE. Reliability and validity of the Visual Gait Assessment Scale for children with hemiplegic cerebral palsy when used by experienced and inexperienced observers. *Gait Posture*. 2008;27(4):648-652.

50. Araujo P a, Kirkwood RN, Figueiredo EM. Validity and intra- and inter-rater reliability of the Observational Gait Scale for children with spastic cerebral palsy. *Rev Bras Fisioter*. 2009;13:267-273.

51. Rathinam C, Bateman A, Peirson J, Skinner J. Observational gait assessment tools in paediatrics - A systematic review. *Gait Posture*. 2014;40(2):279-285.

52. Toro B, Nester C, Farren P. A review of observational gait assessment in clinical practice. *Physiother Theory Pract*. 2003;19:137-150.

53. Gage JR. Con:Interobserver variability of gait analysis. *J Pediatr Orthop*. 2003;23(6):290-291.

54. Laracca E, Stewart C, Postans N, Roberts A. The effects of surgical lengthening of hamstring muscles in children with cerebral palsy - The consequences of pre-operative muscle length measurement. *Gait Posture*. 2014;39(3):847-851.

55. Healy MT, Schwartz MH, Stout JL, Gage JR, Novacheck TF. Is simultaneous hamstring lengthening necessary when performing distal femoral extension osteotomy and patellar tendon advancement? *Gait Posture*. 2011;33(1):1-5.

56. Cook RE, Schneider I, Hazlewood ME, Hillman SJ, Robb JE. Gait Analysis Alters Decision-Making in Cerebral Palsy. *J Pediatr Orthop*. 2003:292-295.

57. Wren TAL, Otsuka NY, Bowen RE, et al. Outcomes of lower extremity orthopedic surgery in ambulatory children with cerebral palsy with and without gait analysis: Results of a randomized controlled trial. *Gait Posture*. 2013;38(2):236-241.

58. De Luca P, Davis RB, Ounpuu S, Rose S, Sirkin R. Alterations in Surgical Decision Making in Patients with Cerebral Palsy based on Three-Dimensional Gait Analysis. *J Pediatr Orthop*. 1997;17(5):608-614.

59. Chang FM, Seidl AJ, Muthusamy K, Meininger AK, Carollo JJ. Effectiveness of instrumented gait analysis in children with cerebral palsy--comparison of outcomes. *J Pediatr Orthop*. 2006;26(5):612-616.

60. Filho MCDM, Yoshida R, Carvalho WDS, Stein HE, Novo NF. Are the recommendations from three-dimensional gait analysis associated with better postoperative outcomes in patients with cerebral palsy? *Gait Posture*. 2008;28(2):316-322.

61. Skaggs DL, Rethlefsen S a, Kay RM, Dennis SW, Reynolds R a, Tolo VT. Variability in gait analysis interpretation. *J Pediatr Orthop*. 2000;20(6):759-764.

62. Chau T. A review of analytical techniques for gait data. Part 1: Fuzzy, statistical and fractal methods. *Gait Posture*. 2001;13(1):49-66.

63. Schwartz MH, Trost JP, Wervey R a. Measurement and management of errors in quantitative gait data. *Gait Posture*. 2004;20(2):196-203.

64. Schwartz MH, Rozumalski A. The gait deviation index: A new comprehensive index of gait pathology. *Gait Posture*. 2008;28(3):351-357.

65. Baker R, McGinley JL, Schwartz MH, et al. The gait profile score and movement analysis profile. *Gait Posture*. 2009;30(3):265-269.

66. Barton GJ, Hawken MB, Scott MA, Schwartz MH. Movement Deviation Profile: A measure of distance from normality using a self-organizing neural network. *Hum Mov Sci*. 2012;31(2):284-294.

67. Desloovere K, Molenaers G, Feys H, Huenaerts C, Callewaert B, Walle P Van de. Do dynamic and static clinical measurements correlate with gait analysis parameters in children with cerebral palsy? *Gait Posture*. 2006;24(3):302-313.

68. Dreher T, Vegvári D, Wolf SL, et al. Long-term effects after conversion of biarticular to monoarticular muscles compared with musculotendinous lengthening in children with spastic diplegia. *Gait Posture*. 2013;37(3):430-435.

69. Pataky TC, Robinson M a., Vanrenterghem J. Vector field statistical analysis of kinematic and force trajectories. *J Biomech*. 2013;46(14):2394-2401.

70. Wolf S, Loose T, Schablowski M, et al. Automated feature assessment in instrumented gait

analysis. *Gait Posture*. 2006;23(3):331-338.

71. Knudson D. Statistical and reporting errors in applied biomechanics research. In: *Proceedings of XXIII International Symposium on Biomechanics in Sports*. Vol 2. Beijing: China Institute of Sport Science.; :811-814.

72. McLaughlin MJ, Sainani KL. Bonferroni, holm, and hochberg corrections: Fun names, serious changes to P values. *PM R*. 2014;6(6):544-546.

73. Pataky TC, Vanrenterghem J, Robinson MA. The probability of false positives in zero-dimensional analyses of one-dimensional kinematic, force and EMG trajectories. *J Biomech*. 2016;49:1468-1476.

74. Knudson D, Lindsey C. Type I and type II errors in correlations of various sample sizes. *Compr Psychol*. 2014;3(1):811-814.

75. Pataky TC. Generalized n-dimensional biomechanical field analysis using statistical parametric mapping. *J Biomech*. 2010;43(10):1976-1982.

76. Simon S, Deutsch S, Nuzzo R, et al. Genu recurvatum in spastic cerebral palsy. *J Bone Joint Surg Am*. 1978;60-A:882-894.

77. Vaughan CL, O'Malley MJ. A gait nomogram used with fuzzy clustering to monitor functional status of children and young adults with cerebral palsy. *Dev Med Child Neurol*. 2005;47(6):377-383.

78. Schmidt-Rohlfing B, Bergamo F, Williams S, et al. Interpretation of surface EMGs in children with cerebral palsy: An initial study using a fuzzy expert system. *J Orthop Res*. 2006;24(3):438-447.

79. Toro B, Nester CJ, Farren PC. Cluster analysis for the extraction of sagittal gait patterns in children with cerebral palsy. *Gait Posture*. 2007;25(2):157-165.

80. Cioni G, Lodesani M, Pascale R, et al. The term diplegia should be enhanced. Part II: Contribution to validation of the new rehabilitation oriented classification. *Eur J Phys Rehabil Med*. 2008;44(2):203-211.

81. Pascale R, Perazza S, Borelli G, et al. The term diplegia should be enhanced. Part III: inter-observer reliability of the new rehabilitation oriented classification. *Eur J Phys Rehabil Med*. 2008;44(2):213-220.

82. Van Gestel L, De Laet T, Di Lello E, et al. Probabilistic gait classification in children with cerebral palsy: A Bayesian approach. *Res Dev Disabil*. 2011;32(6):2542-2552.

83. Bonnefoy-Mazure A, Sagawa YJ, Lascombes P, De Coulon G, Armand S. Identification of gait patterns in individuals with cerebral palsy using multiple correspondence analysis. *Res Dev Disabil*. 2013;34(9):2684-2693.

84. Nieuwenhuys A, Õunpuu S, Van Campenhout A, et al. Identification of joint patterns during gait in children with cerebral palsy: A Delphi consensus study. *Dev Med Child Neurol*. 2016;58(3):306-313.

85. Davids JR, Bagley AM. Identification of Common Gait Disruption Patterns in Children With Cerebral Palsy. *J Am Acad Orthop Surg*. 2014;22(12):782-790.

86. Dobson F, Morris ME, Baker R, Graham HK. Unilateral cerebral palsy: a population-based study of gait and motor function. *Dev Med Child Neurol*. 2011;53(5):429-435.

87. Dobson F, Morris M, Baker R, Wolfe R, Graham H. Clinician agreement on gait pattern ratings in children with spastic hemiplegia. *Dev Med Child Neurol*. 2006;48(6):429-435

88. Riad J, Haglund-Akerlind Y, Miller F. Classification of Spastic Hemiplegic Cerebral Palsy in Children. *J Pediatr Orthop*. 2007;27(7):758-764.

89. de Morais Filho MC, Kawamura CMC, Lopes JJAF, Neves DLD, Cardoso MM de O, Caiafa JJB. Most frequent gait patterns in diplegic spastic cerebral palsy. *Acta Orthop Bras [online]*. 2014;22(4):197-201.

90. Stott NS, Atherton WG, Mackey AH, Galley IJ, Nicol RO, Walsh SJ. The reliability and validity of assessment of sagittal plane deviations in children who have spastic diplegia. *Arch Phys Med Rehabil*. 2005;86(12):2337-2341.

91. Kim DJ, Park ES, Sim EG, Kim KJ, Kim YU, Rha D. Reliability of Visual Classification of Sagittal Gait Patterns in Patients with Bilateral Spastic Cerebral Palsy. *Ann Rehabil Med*. 2011;35(3):354.

92. Xu G, Zhang Y, Begg R. Mining gait pattern for clinical locomotion diagnosis based on clustering techniques. In: Li, X and Zaiane, OR and Li Z, ed. *Advanced Data Mining and Applications Proceedings (ADMA)*. Vol 4093. Lecture Notes in Artificial Intelligence. ; 2006:296-307.

93. Benedetti MG, D'Apote G, Faccioli S, Costi S, Ferrari A. Equinus foot classification in cerebral palsy: An agreement study between clinical and gait analysis assessment. *Eur J Phys Rehabil Med*. 2011;47(2):213-221.

94. Rozumalski A, Schwartz MH. Crouch gait patterns defined using k-means cluster analysis are related to underlying clinical pathology. *Gait Posture*. 2009;30(2):155-160.

95. Krzak JJ, Corcos DM, Damiano DL, et al. Kinematic foot types in youth with equinovarus secondary to hemiplegia. *Gait Posture*. 2015;41(2):402-408.

96. Agostini V, Nascimbeni A, Gaffuri A, Knaflitz M. Multiple gait patterns within the same Winters class in children with hemiplegic cerebral palsy. *Clin Biomech*. 2015;30(9):908-914.

97. McDowell BC, Kerr C, Kelly C, Salazar J, Cosgrove A. The validity of an existing gait classification system when applied to a representative population of children with hemiplegia. *Gait Posture*. 2008;28(3):442-447.

98. Padilla UR. Fuzzy Classification of Hemiplegic Gait Using Kinematic Indicators in Knee. In: Braidot, A and Hadad A, ed. *VI Latin American Congress on Biomedical Engineering (CLAIB 2014)*. Vol 49. IFMBE Proceedings. ; 2014:596-599.

99. Sangeux M, Rodda J, Graham HK. Sagittal gait patterns in cerebral palsy: The plantarflexor-knee extension couple index. *Gait Posture*. 2015;41(2):586-591.

100. Carriero A, Zavatsky A, Stebbins J, Theologis T, Shefelbine SJ. Determination of gait patterns in children with spastic diplegic cerebral palsy using principal components. *Gait Posture*. 2009;29(1):71-75.

101. Ferrari A, Alboresi S, Muzzini S, Pascale R, Perazza S, Cioni G. The term diplegia should be enhanced. Part I: a new rehabilitation oriented classification of cerebral palsy. *Eur J Phys Rehabil Med*. 2008;44(2):195-201.

102. Williams SE, Gibbs S, Meadows CB, Abboud RJ. Classification of the reduced vertical component of the ground reaction force in late stance in cerebral palsy gait. *Gait Posture*. 2011;34(3):370-373.

103. Simon A-L, Ilharreborde B, Megrot F, et al. A Descriptive Study of Lower Limb Torsional Kinematic Profiles in Children With Spastic Diplegia. *J Pediatr Orthop*. 2015;35(6):576-582.

104. Mokkink LB, Terwee CB, Patrick DL, et al. The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes. *J Clin Epidemiol*. 2010;63(7):737-745.

105. Mokkink LB, Terwee CB, Patrick DL, et al. The COSMIN checklist for assessing the methodological quality of studies on measurement properties of health status measurement

instruments: An international Delphi study. *Qual Life Res*. 2010;19(4):539-549.

106. Moher D, Liberati A, Tetzlaff J, Altman DG, Grp P. Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement (Reprinted from Annals of Internal Medicine). *Phys Ther*. 2009;89(9):873-880.

107. Williams SE, Gibbs S, Meadows CB, et al. Classification of the reduced vertical component of the ground reaction force in late stance in cerebral palsy gait. *Gait Posture*. 2011;34(3):370-373.

108. Kamruzzaman J, Begg RK. Support vector machines and other pattern recognition approaches to the diagnosis of cerebral palsy gait. *IEEE Trans Biomed Eng*. 2006;53(12):2479-2490.

109. Zhang B, Zhang Y, Begg RK. Gait classification in children with cerebral palsy by Bayesian approach. *PATTERN Recognit*. 2009;42(4, SI):581-586.

110. Winters TF, Gage JR, Hicks R. Gait patterns in spastic hemiplegia in children and young adults. *J Bone Joint Surg Am*. 1987;69:437-441.

111. Sutherland DH, Davids JR. Common Gait Abnormalities of the Knee in Cerebral Palsy. *Clin Orthop Relat Res*. 1993;288:139-147.

112. Rodda JM, Graham HK, Carson L, Galea MP, Wolfe R. Sagittal gait patterns in spastic diplegia. *J Bone Joint Surg Br*. 2004;86(2):251-258.

113. Carollo J, He Q, Debrunner C. Gait pattern classification in children with cerebral palsy: results from a 12-state hidden Markov model created with a large training set. *Gait Posture*. 2004;20:S1-S59.

114. Zwick EB, Leistritz L, Milleit B, et al. Classification of equinus in ambulatory children with cerebral palsy - Discrimination between dynamic tightness and fixed contracture. *Gait Posture*. 2004;20(3):273-279.

115. Dallmeijer AJ, Baker R, Dodd KJ, Taylor NF. Association between isometric muscle strength and gait joint kinetics in adolescents and young adults with cerebral palsy. *Gait Posture*. 2011;33(3):326-332.

116. Shin H-I, Sung KH, Chung CY, et al. Relationships between Isometric Muscle Strength, Gait Parameters, and Gross Motor Function Measure in Patients with Cerebral Palsy. *YONSEI Med J*. 2016;57(1):217-224.

117. Van Campenhout A, Bar-On L, Aertbeliën E, Huenaerts C, Molenaers G, Desloovere K. Can we unmask features of spasticity during gait in children with cerebral palsy by increasing their walking velocity? *Gait Posture*. 2014;39(3):953-957.

118. O'Malley MJ, Abel MF, Damiano DL, Vaughan CL. Fuzzy clustering of children with cerebral palsy based on temporal-distance gait parameters. *IEEE Trans Rehabil Eng*. 1997;5(4):300-309.

119. Baker R, McGinley JL, Schwartz M, Thomason P, Rodda J, Graham HK. The minimal clinically important difference for the Gait Profile Score. *Gait Posture*. 2012;35(4):612-615.

120. Daniels L, Worthingham C. *Muscle Testing Techniques of Manual Examination*. Philadelphia, PA: W B Saunders; 1986.

121. Mutlu A, Livanelioglu A, Gunel MK. Reliability of Ashworth and Modified Ashworth scales in children with spastic cerebral palsy. *BMC Musculoskelet Disord*. 2008;9:44.

Chapter 2

Literature review and comparison of two statistical methods to evaluate the effect of botulinum toxin treatment on gait in children with cerebral palsy

Angela Nieuwenhuys
Eirini Papageorgiou
Todd Pataky
Tinne De Laet
Guy Molenaers
Kaat Desloovere

Adapted from PlosOne 2016;11(3);1-17

Abstract

Aim

This study aimed at comparing two statistical approaches to analyze the effect of Botulinum Toxin A (BTX-A) treatment on gait in children with a diagnosis of spastic cerebral palsy (CP), based on three-dimensional gait analysis (3DGA) data. Through a literature review, the available expert knowledge on gait changes after BTX-A treatment in children with CP was summarized.

Methods

Part 1 - Intervention studies on BTX-A treatment in children with CP between 4-18 years that used 3DGA data as an outcome measure and were written in English, were identified through a broad systematic literature search. Reported kinematic and kinetic gait features were extracted from the identified studies. **Part 2** - A retrospective sample of 53 children with CP (6.1 ± 2.3years, GMFCS I-III) received 3DGA before and after multilevel BTX-A injections. The effect of BTX-A on gait was interpreted by comparing the results of paired samples t-tests on the kinematic gait features that were identified from literature, to the results of statistical parametric mapping analysis on the kinematic waveforms of the lower limb joints.

Results

Part 1 - 53 kinematic and 33 kinetic features were described in literature. Overall, there is no consensus on which features should be evaluated after BTX-A treatment, as 49 features were reported only once or twice. **Part 2** - Post-BTX-A, both statistical approaches found increased ankle dorsiflexion throughout the gait cycle and increased external rotation of the foot progression angle during stance phase. Statistical parametric mapping analyses additionally found increased knee extension during terminal stance and increased external rotation of the foot progression angle during swing phase.

Conclusion

This study confirms that BTX-A injections are a valuable treatment option to improve gait function in children with CP. However, different statistical approaches may lead to different interpretations of treatment outcome. It is suggested that a clear, definite hypothesis should be stated a priori and a commensurate statistical approach should accompany this hypothesis.

Introduction

Pathological gait is one of the most striking characteristics in children with cerebral palsy (CP)[1]. When spasticity, weakness or other CP-related motor impairments manifest during walking, they may significantly restrict patients at the level of 'activities' and 'participation' of the International Classification of Functioning, Disability and Health[2]. There is a wide variety of treatments which can be applied to improve gait, ranging from conservative treatments such as physiotherapy, orthotics, and Botulinum toxin A (BTX-A) injections to surgical interventions such as single event multilevel surgery and selective dorsal rhizotomy. Depending on a patient's symptoms and age, different treatment modalities might be appropriate.

Gait changes after treatment are often objectively quantified by using three-dimensional gait analysis (3DGA). 3DGA provides a large amount of multivariate kinematic and kinetic waveforms, which have proven to be highly valuable during the clinical decision-making process[3,4]. Researchers and clinicians are challenged to extract and analyze the clinically relevant information from this large amount of data. In general, specific, directed hypotheses are not stated prior to data collection. For example, a typical null hypothesis may state that there are no differences between the knee and ankle kinematics of children with CP pre- and post-treatment. As a consequence, the full gait cycles of the knee and ankle joints should ideally be considered in statistical analysis because the hypothesis implicitly pertains to the full gait cycle and to all knee/ankle kinematic variables. In the literature however, several intervention studies that have examined the effect of treatment on gait, have reduced the amount of 3DGA data a priori by analyzing a number of specific kinematic or kinetic gait features, which refer to a specific time instant of a gait cycle, e.g. peak values[5–9]. The rationale behind the selection of these features is often unclear because those specific variables rarely appear explicitly in hypotheses. Most likely, they were chosen based on available clinical expert knowledge, through literature search, or potentially after viewing the data. Depending on the available clinical expertise, reducing this large amount of data a priori or post-hoc may introduce bias and potential clinically relevant information could be overlooked. Furthermore, by conducting statistical testing on multiple dependent gait features in relatively small sample sizes, the risk of detecting a false positive outcome (Type I error)

increases. In turn, a Bonferroni correction, which is often applied to deal with this risk, will increase the probability of obtaining a false negative result (TypeII error)[10].

In the past years, a promising approach called statistical parametric mapping (SPM) was introduced to the field of biomechanics[11,12]. SPM is a statistical method able to perform hypothesis testing on kinematic and kinetic data in a continuous manner, thereby making a priori data reduction for non-directed hypotheses redundant. It also takes into account the dependency between different time instances of the gait cycle[11,12]. SPM has already been used for example to evaluate whether electromyography time-series of four lower limb muscles during the stance phase are different between children and adults[13]. So far, it has not yet been used to evaluate the outcome of treatment in CP.

After a thorough search of the available literature, following the inclusion criteria described in the methods section of this paper, 223 peer reviewed scientific papers that evaluated the effect of treatment on gait in children with CP based on the analysis of kinematic or kinetic gait features, were identified. Some of the gait features that were analyzed in these papers to quantify the effect of treatment on gait were recurrent in many studies, while some others were only reported a small number of times. Furthermore, it appeared that the definitions of several features were somewhat unclear, making it difficult for researchers to reproduce or confirm the results. An example is the feature 'hip extension during terminal stance'. It is not clear whether this feature refers to a specific time instance of the gait cycle, or whether it refers to the mean value or peak value of the hip during a phase of gait. Furthermore, the phase 'terminal stance' could potentially be defined differently across various studies because patients presenting with a pathological gait might not have a typical 60/40 ratio for stance and swing phase or might not display a typical heel rise in case they do not reach a flat foot position[14,15].

To assess the value of SPM analysis as an alternative approach to feature analysis with regard to the interpretation of treatment outcome based on 3DGA in children with CP, this study focused on one treatment modality, namely BTX-A injections. BTX-A is used to treat spasticity and has been proven to improve function and delay deterioration towards fixed muscle contractures or bony deformities[16,17]. After tone reduction, 3DGA can highlight a child's ability to alter their gait pattern. It can also detect to what extent other clinical motor symptoms such as postural instability or muscle weakness may contribute to the pathological gait pattern[18].

The aim of this study was twofold. First, the available expert knowledge on gait pattern changes in children with CP after BTX-A treatment was summarized. By performing a systematic literature search, an overview of gait features that have been frequently reported in literature and that have been shown to be responsive to BTX-A treatment was created. Secondly, the effect of multi-level BTX-A treatment on gait in a retrospective sample of children with CP was evaluated by comparing the results of the frequently reported feature analysis to the results of SPM analyses on the kinematic waveforms of the lower limb joints. It was expected that SPM would be judged as a valuable alternative statistical approach to describe the effect of BTX-A on gait in a wider and more unbiased perspective than the traditional feature analysis.

Material and methods

Ethical approval for this project was granted by the Medical Ethical Committee of University Hospitals Leuven, reference s56036. All patient information was anonymized prior to statistical analysis. Two major methodological parts were related to the main study goals. The first part involved a literature search to define and select reported gait features that quantify the effect of BTX-A treatment. The second part was an experimental outcome study, comparing SPM analysis to feature analysis in order to interpret the outcome of 3DGA pre- and post-BTX-A treatment in a group of children with CP.

Literature search

A broad systematic search in the databases of Pubmed, Embase, Cinahl, and Web Of Science was performed. Key words included cerebral palsy, diplegia, hemiplegia, quadriplegia, gait analysis, locomotion, walking, gait, feature, parameter, variable, and characteristic. Relevant wildcard symbols were used to ensure all key word variations were searched and if possible, searches were limited to human studies, age, and language. After removal of duplicates, references were screened based on title and abstract. Eligibility of full-text papers was then assessed based on the following inclusion criteria: (a) intervention studies evaluating the effect of treatment on gait, using instrumented 3DGA; (b) the experimental population consisting for at least 80% out of children between the ages of 4-18 years with a diagnosis of the spastic type of CP; (c) a definition of kinematic and/or kinetic features, including at least joint angles or moments; (d) English full text availability in a Belgian library or at request to the author. The following were excluded: case series, literature reviews, and intervention

studies that only reported on indices (such as GPS, GGI, etc.), electromyography features, spatio-temporal parameters or children with dystonia.

The literature search explored all papers that have reported any type of treatment modality to improve gait in children with CP, using 3DGA as an outcome measurement tool. For the purpose of this study, only intervention studies evaluating BTX-A treatment were selected. All kinematic and kinetic features of the papers identified during the review, were extracted. Subsequently, features with different terminology were grouped together in case they had a similar meaning (e.g. minimal hip angle in sagittal plane during stance and maximum hip extension during the gait cycle). Apart from the initial literature search, two reviewers completed each step of the review process and a third reviewer was consulted in case of disagreement. The first search was conducted on December 9, 2013 and it was updated on October 28, 2015.

Experimental outcome study

Patients and treatment characteristics

Patients were retrospectively selected from the database of the Clinical Motion Analysis Laboratory of University Hospital Pellenberg, Leuven, Belgium. We considered children with CP who attended the hospital for BTX-A treatment between 2004 and 2014. Children eligible for this study met the following inclusion criteria: (a) age between 4-18 years, (b) a predominantly spastic diagnosis of CP, (c) walking with or without assistance of aids (GMFCS I-III), (d) BTX-A injections had occurred in at least hamstrings and gastrocnemius muscles (e) a maximum of three months between the date of BTX-A injection and the pre-3DGA, (f) at least 1 month and maximally 4 months between the date of BTX-A injection and the post-3DGA. Exclusion criteria were symptoms of dystonia or ataxia and previous orthopedic surgery.

In case multiple gait analysis sessions were available for one patient, a preference was given to the gait analyses that were collected before and after the first or second BTX-A treatment, because patients are likely to improve more after the first treatments[19]. All BTX-A injections were performed by a pediatric orthopedic surgeon and were part of an integrated multilevel treatment approach which was previously described by Molenaers et al.[20]. BTX-A injections were administered under general (mask) anesthesia, applying a dilution of 100 units (U) of Botox® (Allergan, Inc. Irvine, USA) in 5ml of saline. The muscles selected for treatment

were injected at multiple sites with a maximum of 50U per site and a minimal distance of 5 cm between injection sites. In accordance with the integrated approach, serial stretching casts, an increased number of physiotherapy sessions, and increased use of orthoses were planned after BTX-A treatment.

Data collection procedure

All gait analyses were conducted in the clinical motion analysis laboratory of the University Hospital Pellenberg, using a Vicon system (Oxford Metrics, Oxford, UK) and two AMTI force plates (Advanced Mechanical Technology Inc., Watertown, MA, USA). Children were asked to walk barefoot on a 10m walkway at a self-selected, comfortable speed. Ten to fifteen infra-red VICON cameras captured the position of retroflective markers, which were placed on the bony landmarks of the child according to the Plug-In-Gait marker model of Vicon (Oxford Metrics, Oxford, UK). Nexus software was used to define gait cycles and estimate joint angles over the three anatomical planes. For children with bilateral CP, both legs were eligible for analysis and for children with unilateral CP, only the affected leg was included. Per included leg, the average of two gait trials was analyzed. After a thorough quality check of all included gait trials, the kinematic parameters reported in literature as well as waveform data were exported using custom-made Matlab software (Mathworks, Inc. version 2014a).

Statistical analysis

Only the kinematic joint angle features that were identified during the literature review were considered in the analysis, because good quality kinetic data were not available for all participants. In case the definition of a feature in literature was unclear and we were unable to recalculate it for our own data, it was excluded from the analysis. For each selected feature, a paired samples t-test compared the included group of CP patients pre-treatment versus post-treatment using SPSS Statistics for Windows, version 20.0 (Armonk, NY). The overall probability of making a TypeI error was maintained at α=0.05, by adjusting p-values according to the Holm procedure (a stepwise Bonferroni correction)[21–23].

SPM was performed on the time-normalized gait cycles of the lower limb joint angles across the three anatomical planes, taking into consideration the dependency of all points of each gait cycle. A conservative Bonferroni correction of $\alpha = 0.01$ across the five joints (pelvis, hip, knee, ankle, and foot) maintained a family-wise TypeI error rate of 5%. First, an SPM two-tailed paired t-test compared the mean joint angles of the knee and ankle in the sagittal plane

as well as the foot progression angle before and after treatment. A statistical parametric map or SPM{t} was computed, representing the traditional univariate t-statistic calculated at each point of the gait cycle; this approach is termed 'mass univariate'. Afterwards, Random Field Theory was used to calculate the critical threshold *t* above which only 1% (α=0.01) of equally smooth random data samples' SPM{t} waveforms would be expected to cross[24]. Whenever the experimentally observed SPM{t} exceeded this critical threshold, a supra-threshold cluster probability was computed, which indicated a significant statistical difference between the pre-treatment and post-treatment analyses in that part of the gait cycle. A p-value for each supra-threshold cluster was calculated to specify the probability of discovering a cluster with identical temporal breadth when equally smooth random data would be analyzed[24]. Secondly, an SPM paired Hotellings T^2 was computed to compare the mean joint angles of the pelvis and hip joint before and after BTX-A treatment. The SPM paired Hotellings T^2 statistic takes into account the time dependency of all points of the gait cycle, as well as the covariance of joint kinematics over the three anatomical planes[25]. Post-hoc t-tests were performed if the Hotellings T^2 test presented a statistically significant outcome, i.e. when the critical threshold *t* (α=0.01) was exceeded. These post-hoc t-tests constituted a separate analysis of the pelvis or hip kinematics in the three planes. A full description of this workflow is described in Robinson et al.[13]. All analyses were performed in Python (Python 2.7.2; Enthought Python Distribution, Austin, TX), using open-source SPM1D code (v.0.3; www.spm1D.org).

Results

Literature search

The literature search yielded a total of 2531 titles and abstracts, which was reduced to a selection of 26 papers that evaluated BTX-A treatment using 3DGA in children with CP (Fig 1)[5,6,9,18,26–47]. Fifteen papers reported the effect of BTX-A treatment to the gastrocnemius muscle, sometimes in combination with the soleus muscle[5,6,26–28,30,31,35,37–39,41–43,46]. Besides BTX-A injections to the gastrocnemius, nine papers also included the hamstrings in a multi-level treatment[9,18,29,33,34,36,40,44,47]. Two papers focused solely on BTX-A injections to the hamstrings[32,45]. The papers reported a median of five features (range 2-49). After features with a similar meaning were grouped together, 53 kinematic features (S1-S5 Table) and 33 kinetic features (S6-S8 Table) were identified. Eleven kinematic features, which were

ambiguously defined, could not be included in the statistical analysis of the experimental outcome study (part 2). Figs 2-6 list all 42 features that were selected for statistical analysis in the experimental outcome study and show for each of those features the number of papers that have reported it to be responsive to BTX-A treatment in children with CP.

In general, results were mixed. Almost half of all kinematic features were reported only once (n=12) or twice (n=12). On the other hand, the maximal dorsiflexion angle during stance was reported in 23 papers and 21 of them found an improved maximal dorsiflexion angle post-BTX-A[5,6,9,18,26,27,29–31,34–44,47]. There is also consensus that dorsiflexion in the ankle during the swing phase improves after BTX-A treatment of the gastrocnemius and/or soleus muscle. Seven papers reported an increased maximal dorsiflexion angle during swing[26,29,36,39,41–43] and four papers reported an increased dorsiflexion angle at 50% of the gait cycle[9,18,30,34]. At the level of the knee, seven out of eight papers, which reported the maximal knee flexion angle during swing, agreed that it was not influenced by BTX-A treatment[6,9,18,31–33,40]. Knee flexion angle at initial contact and maximal knee extension angle during stance were reported eight[5,9,18,31–33,40,47] and twelve[5,6,18,27,31–33,37,40,42,44,45] times respectively, yet results are contradicting. Only three out of eight papers reported a decreased knee flexion angle at initial contact[5,32,47] and five out of twelve papers reported an improved knee extension angle during stance[18,27,32,40,45]. In the hip joint, six papers reported no significant effect of BTX-A injections to the hamstrings and/or gastrocnemius muscles on the maximal hip flexion angle during swing[9,18,32,33,40,45].

Of the 33 kinetic features that were identified, 25 were reported only once (n=17) or twice (n=8). The features reported most often are the second peak in the ankle moment curve (during the second half of stance phase)[9,18,28,34,47] and the peak ankle power generation during the gait cycle (maximum positive ankle power)[5,6,26,34,39,44,47]. Three out of five studies agreed that the second peak internal plantarflexion moment increased post-BTX-A[9,18,47]. Six out of seven studies that reported the maximal ankle power generation did not change significantly after BTX-A treatment[5,6,26,34,39,47].

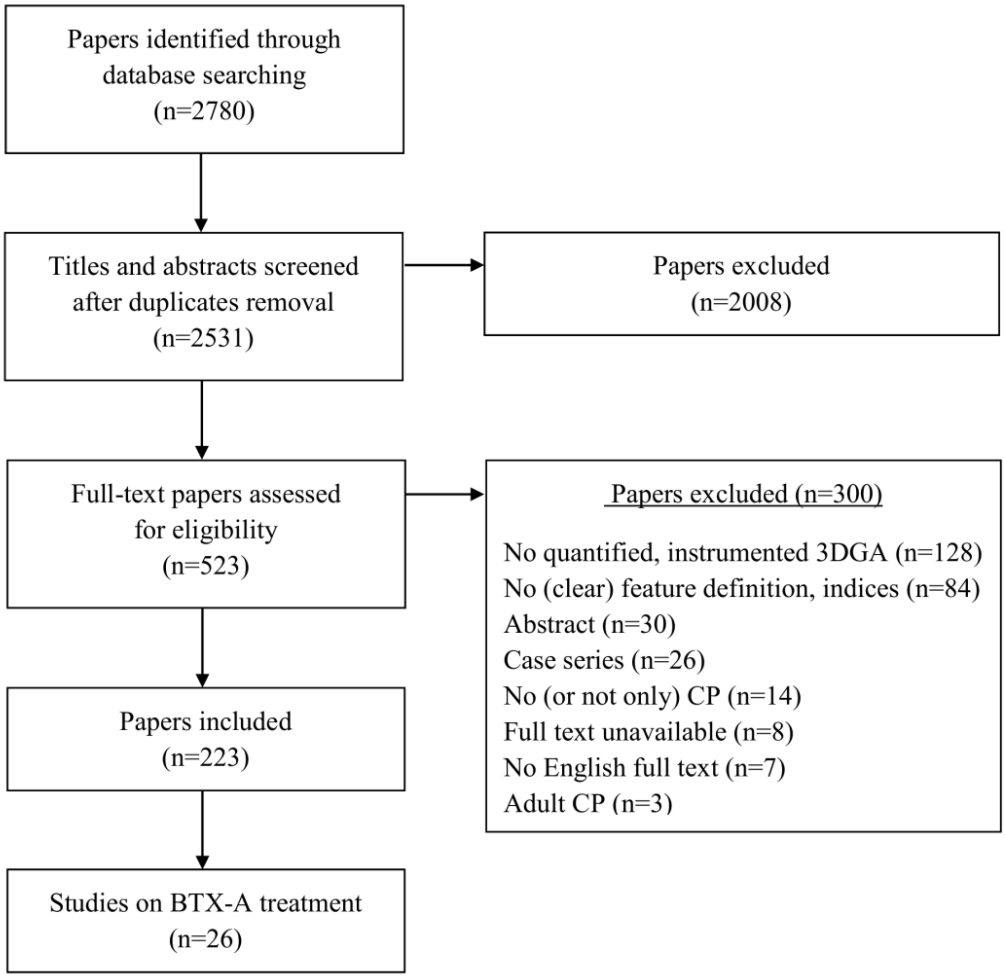
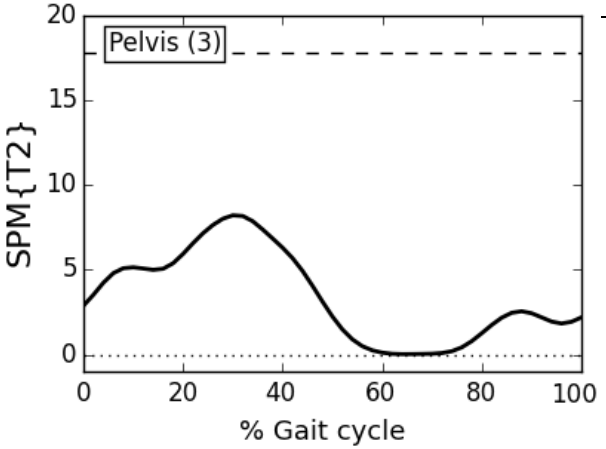


Figure 1. Workflow literature review. This figure describes the workflow which was followed to identify the 26 papers that were included in the literature review. Based on title and abstract, 2008 papers were excluded. After assessing the full-texts, 300 additional papers were excluded using the a priori defined inclusion criteria. In the end, 223 papers that reported on the outcome of treatment in children with CP by means of 3DGA evaluations were identified. Of those 223 papers, 26 reported on the outcome of BTX-A treatment and were included in this study.

(A) SPM analysis



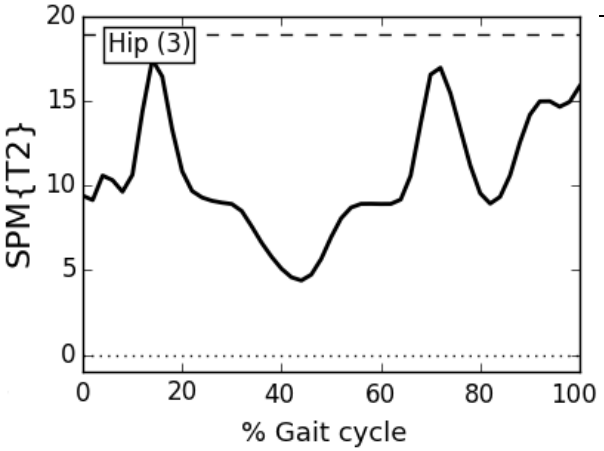
(B) Feature analysis

	Pre BTX-A	Post BTX-A	*p*	(C) Literature Res/Rep(N)
Sagittal plane				
Max. anterior tilt	22.5 *(7)*	22.7 *(7.9)*		0/1
Mean tilt	18.4 *(6.3)*	19 *(7.1)*		3/5
Range of motion	8.9 *(3.7)*	8.2 *(2.8)*		0/4
Coronal plane				
Mean obliquity	-0.8 *(3)*	-1.2 *(3.1)*		0/2
Range of motion	10.5 *(4)*	10.3 *(4.1)*		0/3
Transverse plane				
Mean rotation	-1.8 *(6.2)*	-1.6 *(6.1)*		0/2
Range of motion	19.4 *(5.9)*	18.7 *(6)*		1/3

Figure 2. Pelvis across anatomical planes: Mean (°) and *(SD(°))* of kinematic gait features pre- and post-BTX-A treatment (N=73) compared to SPM analysis (N=73) and findings from literature review.

Panel (A) shows the SPM $\{T^2\}$ statistic ($\alpha = 0.01$) as a function of the gait cycle. The critical threshold (wide dashes) was not exceeded, indicating no significant improvement of BTX-A treatment on the pelvic joint kinematics across the three anatomical planes. Panel (B) shows the mean (°) and (SD) of features extracted from literature. No significant differences were found between pre- and post-BTX-A treatment based on Holm's adjusted p-value (all $p > 0.05$); Max. = maximum. Panel (C) indicates the results from literature review. Res/Rep shows the number of papers that reported the feature to be responsive to BTX-A / number of papers that reported the feature.

(A) SPM analysis



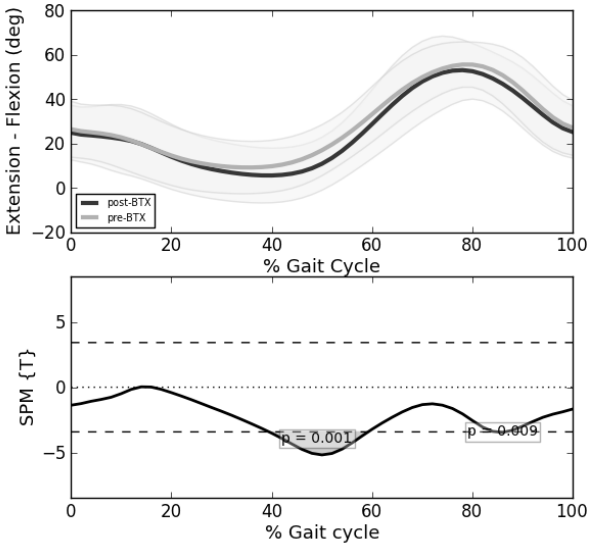
(B) Feature analysis

	Pre BTX-A	Post BTX-A	*p*	(C) Literature Res/Rep(N)
Sagittal plane				
Angle at initial contact	38.7 *(9.3)*	39.1 *(10.1)*		1/5
Max. extension during ST	-1.6 *(9)*	-3.4 *(8.2)*		3/5
Max. flexion during SW	46.8 *(9.5)*	45.1*(9.7)*		0/6
Range of motion during ST	40.3 *(11.5)*	42.7 *(9.9)*		0/2
Range of motion	48.5 *(11.9)*	48.8 *(10)*		2/3
Coronal plane				
Mean angle during ST	3.7 *(3.7)*	2.8 *(4.4)*		0/3
Mean angle during SW	-2.8 *(3.3)*	-3.8 *(3.9)*		1/3
Max. abduction angle	-6.6 *(3.7)*	-7.3 *(4.1)*		0/1
Max. adduction angle	8.1 *(4)*	6.9 *(4.4)*		0/1
Range of motion	14.6 *(3.6)*	14.2 *(3.7)*		0/2
Transverse plane				
Angle at initial contact	2.1 *(10)*	-0.6 *(10.6)*		1/3
Angle at 50% of ST phase	6.1 *(9.4)*	4.7 *(9.2)*		0/2
Angle at toe-off	7.5 *(9.6)*	6.8 *(10.2)*		1/1
Angle at 50% of SW phase	5.2 *(9.4)*	3.6 *(9.6)*		1/3
Mean rotation during ST	6.4 *(9.1)*	5 *(9.1)*		1/1

Figure 3. Hip across anatomical planes: Mean (°) and *(SD(°))* of kinematic gait features pre- and post-BTX-A treatment (N=73) compared to SPM analysis (N=73) and findings from literature review.

Panel (A) shows the SPM $\{T^2\}$ statistic ($\alpha = 0.01$) as a function of the gait cycle. The critical threshold (wide dashes) was not exceeded, indicating no significant improvement of BTX-A treatment on the hip joint kinematics across the three anatomical planes. Panel (B) shows the mean (°) and (SD) of features extracted from literature. No significant differences were found between pre- and post-BTX-A treatment based on Holm's adjusted p-value (all $p > 0.05$); ST = stance; SW = swing; Max. = maximum. Panel (C) indicates the results from literature review. Res/Rep shows the number of papers that reported the feature to be responsive to BTX-A / number of papers that reported the feature.

(A) SPM analysis



(B) Feature analysis

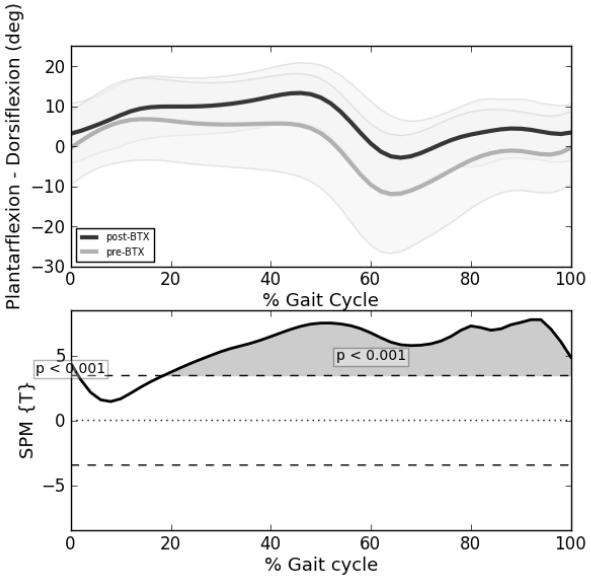
	Pre BTX-A	Post BTX-A	*p*	Res/Rep(N)
Sagittal plane				
Angle at initial contact	26.4 *(12.6)*	24.8 *(12.2)*		3/8
Max. flexion during ST	36.1 *(10.1)*	34.7 *(10.6)*		0/4
Max. extension during ST	7.1 *(11.9)*	4.1 *(12.4)*		5/12
Angle at toe-off	32.3 *(9.5)*	31.9 *(11.5)*		0/1
Max. flexion during SW	60.8 *(10)*	58.4 *(11.1)*		1/8
Range of motion during ST	29 *(8.6)*	30.6 *(7.3)*		1/2
Range of motion	53.7 *(15)*	54.3 *(12.8)*		2/5

(C) Literature

Figure 4. Knee in sagittal plane: Mean (°) and *(SD(°))* of kinematic gait features pre- and post-BTX-A treatment (N=73) compared to SPM analysis (N=73) and findings from literature review.

Panel (A) shows two graphs. The top graph shows the mean kinematics of the knee in the sagittal plane of 73 included legs pre-BTX-A treatment (light grey) versus post-BTX-A treatment (dark gray). The bottom graph represents the SPM {T} statistic ($\alpha = 0.01$) as a function of the gait cycle. The critical threshold t=3.425 (wide dashes) was exceeded at 41-59% and at 86% of the gait cycle, indicating a significant improvement of BTX-A treatment on the knee joint kinematics in the sagittal plane. Panel (B) shows the mean (°) and (SD) of features extracted from literature. No significant differences were found between pre- and post-BTX-A treatment based on Holm's adjusted p-value (all $p > 0.05$); ST = stance; SW = swing; Max. = maximum. Panel (C) indicates the results from literature review. Res/Rep shows the number of papers that reported the feature to be responsive to BTX-A / number of papers that reported the feature.

(A) SPM analysis



(B) Feature analysis

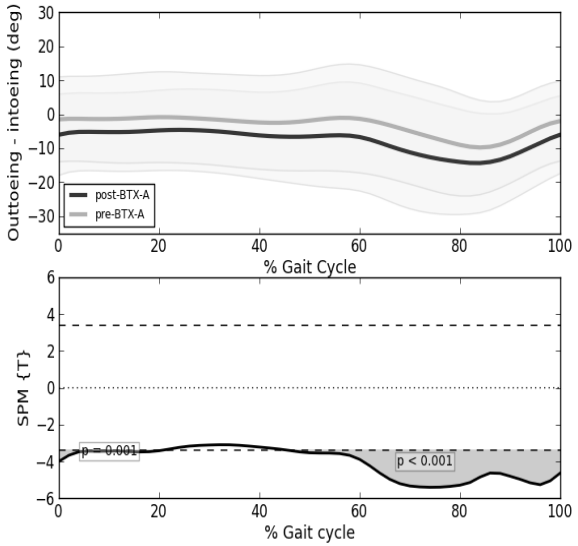
	Pre BTX-A	Post BTX-A	*p*	Res/Rep(N)
Sagittal plane				
Angle at initial contact	-0.4 *(9.2)*	3.1 *(7.5)*	*	9/13
Max. angle between 0-25% GC	7.4 *(10.2)*	9.1 *(7.8)*		3/4
Max. angle during ST	10.5 *(10.2)*	15.3 *(7.1)*	*	21/23
Angle at 50% of ST phase	5.7 *(10.5)*	10.5 *(5.9)*	*	1/1
Max. dorsiflexion during SW	2 *(9)*	6.7 *(6.4)*	*	7/7
Max. plantarflexion during SW	-15.4 *(14)*	-6.8 *(9.5)*	*	1/3
Angle at 50% of SW phase	-2.7 *(11.7)*	4.4 *(6.5)*	*	4/4
Max. plantarflexion angle	-15.5 *(14)*	-6.9 *(9.5)*	*	1/3
Range of motion during SW	17.4 *(9.8)*	13.4 *(6.7)*	*	2/2
Range of motion during push-off	23.3 *(9.3)*	21 *(7.8)*	*	0/3
Range of motion	26.3 *(10.5)*	22.3 *(7.6)*	*	3/5
Timing of max. angle during ST (%GC)	28.9 *(15.3)*	39 *(12.8)*	*	3/3

(C) Literature

Figure 5. Ankle in sagittal plane: Mean (°) and *(SD(°))* of kinematic gait features pre- and post-BTX-A treatment (N=73) compared to SPM analysis (N=73) and findings from literature review.

Panel (A) shows two graphs. The top graph shows the mean kinematics of the ankle in the sagittal plane of 73 included legs pre-BTX-A treatment (light grey) versus post-BTX-A treatment (dark gray). The bottom graph represents the SPM {T} statistic ($\alpha = 0.01$) as a function of the gait cycle. The critical threshold t=3.454 (wide dashes) was exceeded between 0-2% and 22-100% of the gait cycle, indicating a significant improvement of BTX-A treatment on ankle dorsiflexion during the gait cycle. Panel (B) shows the mean (°) and (SD) of features extracted from literature. * indicates a significant difference between pre- and post-BTX-A treatment based on Holm's adjusted p-value ($\alpha < 0.05$); GC = gait cycle; ST = stance; SW = swing; Max. = maximum. Panel (C) indicates the results from literature review. Res/Rep shows the number of papers that reported the feature to be responsive to BTX-A / number of papers that reported the feature.

(A) SPM analysis



(B) Feature analysis

	Pre BTX-A	Post BTX-A	*p*	Res/Rep(N)
Mean foot progression during stance	-1.9 *(13.9)*	-6 *(12)*	*	2/3

(C) Literature

Figure 6. Foot progression angle: Mean (°) and *(SD(°))* of kinematic gait features pre- and post-BTX-A treatment (N=73) compared to SPM analysis (N=73) and findings from literature review.

Panel (A) shows two graphs. The top graph shows the mean kinematics of the foot progression angle of 73 included legs pre-BTX-A treatment (light grey) versus post-BTX-A treatment (dark gray). The bottom graph represents the SPM {T} statistic ($\alpha = 0.01$) as a function of the gait cycle. The critical threshold t=3.390 (wide dashes) was exceeded between 0-22% and 45-100% of the gait cycle, indicating significantly increased outtoeing post-BTX-A treatment. Panel (B) shows the mean (°) and (SD) of features extracted from literature. * indicates a significant difference between pre- and post-BTX-A treatment based on Holm's adjusted p-value ($\alpha < 0.05$). Panel (C) indicates the results from literature review. Res/Rep shows the number of papers that reported the feature to be responsive to BTX-A / number of papers that reported the feature.

Experimental outcome study

A total of 53 patients were included in this study. Patient characteristics are described in Table 1. The majority of patients were diagnosed with bilateral CP (n=36) and GMFCS level I (n=25). There was a median of 26.5 days (range 1-91 days) on average between the pre-3DGA and the date of BTX-A treatment. The median time between the BTX-A treatment session and the post-BTX-A 3DGA was 58 days (range 45-114 days). Seventy-three legs were included in statistical analysis. The median dosage of BTX-A injected into multiple sites of the hamstrings was 4U/kg body weight with a range of 2 to 6 U/kg body weight. A median dosage of 4 U/kg body weight was also administered to the gastrocnemius muscle, spread over different sites, with a range of 2 to 7.5 U/kg body weight. Often iliopsoas, adductors, rectus femoris, soleus, and tibialis posterior were also included in the multilevel BTX-A treatment, though less frequently (Table 2). In accordance to the integrated treatment approach, all children received serial stretching casts for the lower and/or upper legs for a period between 1 to 4 weeks. Children with unilateral CP also received casts for both legs with the aim of obtaining a more symmetric gait pattern.

Table 1. Patient characteristics (N=53).

Gender	
Male	18
Female	35
Diagnosis	
Bilateral CP	36
Unilateral CP	17
Mean age (at time of pre-3DGA) (years, (SD))	6.1 (2.3)
Mean weight (at time of pre-3DGA) (kg, (SD))	20.1 (7.0)
Mean height (at time of pre-3DGA) (cm, (SD))	114.0 (14.6)
GMFCS	
Level I	25
Level II	17
Level III	11
Walking aids during 3DGA	
None	43
Support of one hand	1
Kayewalker	9

3DGA=three-dimensional gait analysis; CP=cerebral palsy.

Table 2. Muscles treated with BTX-A (N=73 treated limbs).

Muscles	**Number of limbs injected**	**Median dose U/kg body weight (range)**
Iliopsoas*	52	2 (1-3)
Adductors	37	1.5 (1-3)
Rectus Femoris	11	1.5 (0.75-2)
Hamstrings	73	4 (2-6)
Gastrocnemius*	73	4 (2-7.5)
Soleus	18	2 (1-3)
Tibialis posterior	5	2 (1.5-2)

* Median dose and range are based on 71 limbs, as dosages were unavailable for two limbs.

Figs 2-6 describe the results for all statistical analyses. For the pelvis and hip joint, neither of the statistical analyses found significant changes. Post-BTX-A treatment, SPM did find significantly improved knee extension between 41-59% and a slightly earlier and lower peak knee flexion at 86% of the gait cycle. The ankle dorsiflexion significantly increased post-BTX-A treatment between 0-2% and 22-100% of the gait cycle. The foot progression angle showed increased outtoeing between 0-22% and 45-100% of the gait cycle. Figs 2-6 also show that out of 42 features, 11 ankle joint features in the sagittal plane and the mean foot progression angle during stance were found to be significantly improved after BTX-A treatment.

Discussion

In this study, the hypothesis was tested that lower limb joint kinematics of children with a spastic diagnosis of CP would improve toward a more typical gait pattern post-BTX-A treatment. Two statistical approaches were compared. On the one hand, kinematic gait features that have previously been reported in literature, were analyzed and on the other hand SPM analyses were conducted. Both approaches detected gait changes, mainly at the ankle. Both feature and SPM analyses concluded that no changes in gait occurred at the level of the pelvis and hip. After treatment, based on SPM analysis, a significantly improved knee extension during stance, an earlier peak knee flexion during swing, and increased ankle dorsiflexion and outtoeing throughout most of the gait cycle were noted. Post-BTX-A, feature analysis also highlighted improved dorsiflexion of the ankle at different time points of the gait

cycle, and additionally an increased outtoeing during stance, but no effect at the knee. From literature, it was shown that results of BTX-A treatment based on feature analyses are generally mixed. Furthermore, feature definitions were not always clear enough to allow us to recalculate them on our own data. For each joint, a more detailed discussion of the presented results compared to literature is presented below.

Pelvis and hip

The SPM and feature analyses of the experimental outcome study did not highlight significant effects of BTX-A treatment on the pelvic and hip kinematics in the three anatomical planes. Along the same lines, few papers in literature report a significant change of pelvic and hip kinematics post-BTX-A treatment. Three out of five studies evaluating 'mean pelvic tilt in the sagittal plane', reported a significantly higher anterior tilt after treatment[32,33,44]. Corry et al.[32] reported an increased anterior tilt as a source of concern after hamstrings injection if the psoas was left untreated. In the present study, 71% of treated limbs also received psoas injections, hence an increased pelvic tilt after BTX-A was not expected. In the hip, literature frequently reported on three kinematic features, namely angle at initial contact, maximal hip extension during stance, and maximal hip flexion during swing. Galli et al.[5] evaluated BTX-A injections to the gastrocnemius muscle and reported a slight deterioration towards an increased hip flexion throughout the gait cycle, which emphasizes the need for a multilevel treatment approach. Depending on whether the hip flexors are included in the multilevel BTX-A treatment, the maximal hip extension in stance may significantly improve or not. This was demonstrated by Desloovere et al.[33], Svehlik et al.[44] and also by Papadonikolakis et al.[40] who have reported an increase in maximal hip extension during stance, but only in the group of patients who received a multi-level BTX-A treatment. However, it must be noted that neither of these studies attempted to maintain the TypeI error rate at 0.05 by accounting for the covariance of multiple dependent gait features, so a false positive result cannot be excluded.

Knee

Compared to the pre-BTX-A condition, SPM analysis of the experimental data noted a significantly improved knee extension and an increased slope towards flexion during terminal stance and pre-swing (between 41% - 59% of the gait cycle). This result may be related to the hamstrings injections. Spasticity of the hamstrings could impede a sufficient amount of knee extension during stance and at the end of swing phase. After BTX-A injection into the hamstrings, a reduction in tone can be achieved; hence improved knee extension during stance

can be expected. In addition, an improvement in knee extension could also be expected after BTX-A injections to the gastrocnemius. On the contrary, no significant differences were found based on the paired samples t-tests on seven gait features of the knee. It is possible that the two features related to terminal stance and pre-swing, namely 'maximal knee extension angle during stance' and 'knee flexion angle at toe-off', are insufficient to characterize these changes in the knee motion during gait post-BTX-A. At the level of the knee, changes in gait might be better characterized by considering the terminal stance and pre-swing phase as a whole. SPM analyses of gait phases are easily interpretable and avoid the need of defining additional features which further contributes to the problem of 'multiple t-testing'. Alternatively, if feature analysis is preferred, we may also conclude that it would be interesting to define an additional feature, for instance knee flexion velocity during pre-swing. In literature, results are mixed as five out of twelve papers found a significant increase in maximal knee extension during stance[18,27,32,40,45]. It should be noted that all papers who did not report an improved knee extension during stance either focused the BTX-A treatments solely on the triceps surae[5,6,31,37,42] or performed the post-3DGA evaluation on average more than one year later[33,44].

SPM analysis also indicated a significant change in a short phase of the knee kinematics during mid-swing, possibly indicating a shift in timing of achieving maximal knee flexion during swing. Unfortunately, the definition of 'amount of delayed knee flexion in swing', which was reported twice in literature with mixed findings, could not be interpreted and reproduced in our analysis[9,18]. The only feature during swing that was analyzed, is 'maximal knee flexion angle during swing', which was found to be unchanged after treatment. In literature, this feature had significantly increased in only one out of eight papers[5]. Galli et al. [5] did not control for an increased TypeI error risk in their analysis of seventeen kinematic and kinetic features. Nevertheless, their study solely investigated treatment of the gastrocnemius muscle, while the peak knee flexion during swing may be predominantly related to rectus femoris spasticity. In the current study, the average peak knee flexion during swing was around 60°, which is well within the one standard deviation around the mean of the typically developing children in the present study. This may explain the fact that the rectus femoris was only injected in 15% of treated limbs. Of the other seven studies that reported no changes in the peak knee flexion during swing, only Papadonikolakis et al.[40] analyzed short-term changes

after BTX-A in a patient group of whom some also received rectus femoris injections. However, it is unclear for how many patients the rectus femoris was included.

Ankle

SPM analysis of the experimental data highlighted a significant increase in dorsiflexion at initial contact after BTX-A treatment. From mid-stance through the remaining part of the gait cycle, the dorsiflexion angle was shown to be significantly increased as well. Features evaluated using paired samples t-tests resembled the results of the SPM analysis. Also in literature, many papers have provided evidence to support these results[5,6,9,18,26,27,29–31,34–44,47]. These results were expected as many CP children suffer from gastrocnemius spasticity potentially leading to equinus gait, which is characterized by an abnormal plantarflexion angle throughout the gait cycle. By means of BTX-A injections, the resulting tone reduction facilitates the motion towards dorsiflexion in stance and to clear the foot during swing phase. In the context of an integrated spasticity treatment, the serial stretching casts which were applied during the first weeks after BTX-A might further support this effect. The study of Bottos et al.[28] is the only paper not reporting a short-term beneficial effect of BTX-A on the ankle joint kinematics. However these results are probably due to a limited sample size (BTX-A group, n=5 and BTX-A plus casting group, n=5).

Foot

At the level of the foot, SPM analysis identified significantly increased outtoeing between 0-22% and 45-100% of the gait cycle. In part, this significant result was confirmed by the only feature which was identified from literature, namely ‘mean foot progression angle during the stance phase’. Compared to literature, one study by Desloovere et al.[33] did not find a significant effect on the foot progression angle after BTX-A treatment. However, the post-3DGA evaluation in this study was on average one year and ten months post-BTX-A injections, making it unlikely to detect a therapeutic effect from the treatment. Two other studies also reported an increased outtoeing after BTX-A[9,18]. They both performed Bonferroni corrections, making a false positive result unlikely. In these two studies, as well as in the presented experimental study, all children received injections to the gastrocnemius along with a period of stretching casts to the lower legs. Hence, more gait improvements can be expected in the distal joints such as the ankle and foot.

Literature

Apart from different statistical approaches and whether or not the Type I error is controlled, another factor that could account for differences in results among the included papers could be the small sample sizes that were included in the studies. With twelve papers evaluating a patient group of less than 20 patients, it is possible that a bias was present in the selected experimental population[5,6,27,28,31,32,34,37,44–47]. The timing of the post-BTX-A 3DGA ranged from two weeks to more than one year across the included papers. This may result in heterogeneous conclusions, given the limited time frame in which BTX-A injections are effective and the importance to acknowledge the effect of other parameters, such as age, on gait when timing of follow-up increases. Naturally, the different clinical characteristics of experimental patient groups as well as the diversity regarding the different muscles which were treated and the dosage of BTX-A, may also contribute to different outcomes. However, a detailed analysis of all these factors would require a thorough evaluation of the methodological quality of the papers, which was beyond the scope of the current study.

Limitations

Limitations of the current study need to be addressed. While creating an overview of the features from literature, rather subjective judgments were made on whether or not particular features could be merged. To avoid bias as much as possible all features were first independently judged by two reviewers, after which a third reviewer was consulted in cases of disagreement. This task was also complicated because a number of features were not clearly defined and thus not easily recalculated. Consequently, it was not possible to include all reported features in the outcome study. Another limitation was the study population, which was recruited from the retrospective database of the University Hospital Pellenberg, Belgium, where every child is treated according to the same rehabilitation guidelines. As a consequence, generalization of results is limited and could be improved through multi-center studies. It was reported by Molenaers et al.[19] that the first two BTX-A treatments are most effective in increasing function. In the intervention study, thirteen patients who had received a third BTX-A treatment and one who had received a fourth treatment were included, which could have led to a reduced effect of BTX-A compared to earlier studies. A sample selection, only containing children after their first or second BTX-A treatment, was not always possible because children were often too young for a 3DGA when they were receiving their first treatments.

Conclusion

In conclusion, both the outcome of the current study as well as literature reports conclude that BTX-A injections are a valuable treatment option to improve gait in children with CP. The effects were mainly observed at the ankle joint and to a lesser extent at the knee. Different results can be explained by various factors which have been illustrated in this study. The key issue which was discussed in this study, is the statistical analysis used to analyze 3DGA data. In literature, the effect on specific gait features is traditionally reported. Considering that over half of all extracted features were only reported once or twice for a list of 26 studies, it can be concluded that there is no consensus on which features should be evaluated to assess the effect of BTX-A on the gait pattern of children with CP. In addition, the risk of obtaining false positive results (Type I error) quickly increases when multiple dependent gait features are analyzed and attempts to control this risk using a Bonferroni correction were in turn decreasing power (increasing the chance of Type II error).

The present study compared this frequently reported feature analysis to SPM. The findings suggest that both statistical methods might be appropriate to analyze kinematic and kinetic data to examine the effect of BTX-A on gait. However, it is suggested that a clear, definite hypothesis should be stated a priori and an adequate, statistical approach should be selected to accompany this hypothesis. When an analysis of features is preferred following a specific hypothesis, it is noted that alternatives to the Bonferroni correction are available to deal with the risk of making a Type I error[10,21,22]. The currently presented literature review could be a guide for feature selection. When reporting features, care should be taken to only include features with an unambiguous definition that is clinically meaningful. For example ‘maximal dorsiflexion during the gait cycle’ could, depending on the clinical presentation of the child, occur during loading response, at the beginning of the third rocker, or during swing phase and might thus not be very meaningful to a clinician. If it is difficult to specifically hypothesize on the direction and magnitude of changes post-treatment, SPM analysis might be preferred. SPM analysis allows the analysis of kinematic and kinetic waveforms as a whole, or particular gait phases, e.g. the swing phase for the knee joint. It has the advantage of making a priori data reduction redundant and also taking into account the covariance of all points of the gait cycle. Nevertheless, at this time, it is not possible to robustly address the co-variation of different joints (e.g. knee and ankle).

For this study, we chose to illustrate the value of SPM by analyzing the effect of BTX-A

treatment on gait because treatment protocols are well standardized and because the effect of this type of treatment has often been reported in literature. Furthermore, the effect of BTX-A is evaluated rather quickly after treatment, ensuring little interference of other treatments or other factors such as age, that may also influence gait. However, it would also be interesting to use SPM to analyze the effects of other treatments in children with CP, such as selective dorsal rhizotomy or orthopedic surgery. SPM might even be more appropriate for pathologies where kinematic and kinetic gait deviations post-treatment have not yet been routinely recognized by clinicians. Non-directed hypotheses evaluated with SPM analysis may help to reduce the wealth of 3DGA data and aid the construction of more specific, directed hypotheses in the future.

Acknowledgements

We thank MSc. Frieke Ostijn and MSc. Jasmine Persoon for their contributions during the systematic search of the literature and for assisting us during the patient selection process.

References

1. Ostensjø S, Carlberg EB, Vøllestad NK. Motor impairments in young children with cerebral palsy: relationship to gross motor function and everyday activities. *Dev Med Child Neurol*. 2004;46(9):580-589.
2. Rosenbaum P, Stewart D. The World Health Organization International Classification of Functioning, Disability, and Health: A Model to Guide Clinical Thinking, Practice and Research in the Field of Cerebral Palsy. *Semin Pediatr Neurol*. 2004;11(1):5-10.
3. Wren TAL, Otsuka NY, Bowen RE, et al. Influence of gait analysis on decision-making for lower extremity orthopaedic surgery: Baseline data from a randomized controlled trial. *Gait Posture*. 2011;34(3):364-369.
4. Cook RE, Schneider I, Hazlewood ME, Hillman SJ, Robb JE. Gait Analysis Alters Decision-Making in Cerebral Palsy. *J Pediatr Orthop*. 2003:292-295.
5. Galli M, Cimolin V, Valente EM, Crivellini M, Ialongo T, Albertini G. Computerized gait analysis of botulinum toxin treatment in children with cerebral palsy. *Disabil Rehabil*. 2007;29(8):659-664.
6. Zwick EB, Svehlik M, Steinwender G, et al. Short-term effects of botulinum toxin A and serial casting on triceps surae muscle length and equinus gait in children with cerebral palsy. *Ces a Slov Neurol a Neurochir*. 2009;72(6):553-558.
7. Dreher T, Vegvári D, Wolf SL, et al. Long-term effects after conversion of biarticular to monoarticular muscles compared with musculotendinous lengthening in children with spastic diplegia. *Gait Posture*. 2013;37(3):430-435.
8. Marconi V, Hachez H, Renders A, Docquier P-L, Detrembleur C. Mechanical work and energy consumption in children with cerebral palsy after single-event multilevel surgery. *Gait Posture*.

2014;40(4):633-639.

9. Molenaers G, Desloovere K, De Cat J, et al. Single event multilevel botulinum toxin type A treatment and surgery: similarities and differences. *Eur J Neurol*. 2001;8 Suppl 5:88-97.
10. Knudson D. Significant and meaningful effects in sports biomechanics research. *Sports Biomech*. 2009;8(1):96-104.
11. Pataky TC. Generalized n-dimensional biomechanical field analysis using statistical parametric mapping. *J Biomech*. 2010;43(10):1976-1982.
12. Pataky TC, Robinson M a., Vanrenterghem J. Vector field statistical analysis of kinematic and force trajectories. *J Biomech*. 2013;46(14):2394-2401.
13. Robinson M a., Vanrenterghem J, Pataky TC. Statistical Parametric Mapping (SPM) for alpha-based statistical analyses of multi-muscle EMG time-series. *J Electromyogr Kinesiol*. 2015;25(1):14-19.
14. Perry J. *Gait Analysis: Normal and Pathological Function*. Thorofare, NJ, Slack; 1992.
15. Baker R. Limitations of the conventional subdivision of the gait cycle. In: *Measuring Walking, a Handbook of Clinical Gait Analysis*. Vol ; 213AD:216-217.
16. Graham HK, Aoki KR, Autti-Rämö I, et al. Recommendations for the use of botulinum toxin type A in the management of cerebral palsy. *Gait Posture*. 2000;11(1):67-79.
17. Lukban MB, Rosales RL, Dressler D. Effectiveness of botulinum toxin A for upper and lower limb spasticity in children with cerebral palsy: a summary of evidence. *J Neural Transm*. 2009;116(3):319-331.
18. Desloovere K, Molenaers G, Jonkers I, et al. A randomized study of combined botulinum toxin type A and casting in the ambulant child with cerebral palsy using objective outcome measures. *Eur J Neurol*. 2001;8 Suppl 5:75-87.
19. Molenaers G, Fagard K, Van Campenhout A, Desloovere K. Botulinum toxin A treatment of the lower extremities in children with cerebral palsy. *J Child Orthop*. 2013;7(5):383-387.
20. Molenaers G, Desloovere K, Eyssen M, De Cat J, Jonkers I, De Cock P. Botulinum toxin type A treatment of cerebral palsy: an integrated approach. *Eur J Neurol*. 6:S51-S57.
21. McLaughlin MJ, Sainani KL. Bonferroni, holm, and hochberg corrections: Fun names, serious changes to P values. *PM R*. 2014;6(6):544-546.
22. Holm S. A Simple Sequentially Rejective Multiple Test Procedure. *Scand J Stat*. 1979;6(2):65-70.
23. J. G. Holm-Bonferroni sequential correction: an excel calculator (1.2). 2013.
24. Adler RJ, Taylor JE. *Random Fields and Geometry*. Vol 1st ed. New York: Springer-Verlag New York; 2007.
25. Pataky TC. One-dimensional statistical parametric mapping in Python. *Comput Methods Biomech Biomed Engin*. 2012;15(3):295-301.
26. Ackman JD, Russman BS, Thomas SS, et al. Comparing botulinum toxin A with casting for treatment of dynamic equinus in children with cerebral palsy. *Dev Med Child Neurol*. 2005;47(9):620-627.
27. Bang MS, Chung SG, Kim SB, Kim S-J. Change of dynamic gastrocnemius and soleus muscle length after block of spastic calf muscle in cerebral palsy. *Am J Phys Med Rehabil*. 2002;81(10):760-764.
28. Bottos M, Benedetti MG, Salucci P, Gasparroni V, Giannini S. Botulinum toxin with and without casting in ambulant children with spastic diplegia: A clinical and functional

assessment. *Dev Med Child Neurol*. 2003;45(11):758-762.

29. Boyd RN, Graham JEA, Nattrass GR, Graham HK. Medium-term response characterisation and risk factor analysis of botulinum toxin type A in the management of spasticity in children with cerebral palsy. *Eur J Neurol*. 1999;6(4):S37-S45.

30. Boyd RN, Pliatsios V, Starr R, Wolfe R, Graham HK. Biomechanical transformation of the gastroc-soleus muscle with botulinum toxin A in children with cerebral palsy. *Dev Med Child Neurol*. 2000;42(1):32-41.

31. Corry IS, Cosgrove AP, Duffy CM, McNeill S, Taylor TC, Graham HK. Botulinum toxin a compared with stretching casts in the treatment of spastic equinus: A randomised prospective trial. *J Pediatr Orthop*. 1998;18(3):304-311.

32. Corry IS, Cosgrove AP, Duffy CM, Taylor TC, Graham HK. Botulinum toxin A in hamstring spasticity. *Gait Posture*. 1999;10(3):206-210.

33. Desloovere K, Molenaers G, De Cat J, et al. Motor function following multilevel botulinum toxin type A treatment in children with cerebral palsy. *Dev Med Child Neurol*. 2007;49(1):56-61.

34. Detrembleur C, Lejeune TM, Renders A, Van Den Bergh PYK. Botulinum toxin and short-term electrical stimulation in the treatment of equinus in cerebral palsy. *Mov Disord*. 2002;17(1):162-169.

35. Galli M, Crivellini M, Santambrogio GC, Fazzi E, Motta F. Short-term effects of "botulinum toxin A" as treatment for children with cerebral palsy: Kinematic and kinetic aspects at the ankle joint. *Funct Neurol*. 2001;16(4):317-323.

36. Kay RM, Rethlefsen SA, Fern-Buneo A, Wren TAL, Skaggs DL. Botulinum toxin as an adjunct to serial casting treatment in children with cerebral palsy. *J Bone Jt Surg - Ser A*. 2004;86(11):2377-2384.

37. Klotz MCM, Wolf SI, Heitzmann D, Gantz S, Braatz F, Dreher T. The influence of botulinum toxin a injections into the calf muscles on genu recurvatum in children with cerebral palsy pediatrics. *Clin Orthop Relat Res*. 2013;471(7):2327-2332.

38. Lee JH, Sung IY, Yoo JY, Park EH, Park SR. Effects of different dilutions of botulinum toxin type A treatment for children with cerebral palsy with spastic ankle plantarflexor: A randomized controlled trial. *J Rehabil Med*. 2009;41(9):740-745.

39. Metaxiotis D, Siebel A, Doederlein L. Repeated botulinum toxin A injections in the treatment of spastic equinus foot. *Clin Orthop Relat Res*. 2002;(394):177-185.

40. Papadonikolakis AS, Vekris MD, Korompilias A V, Kostas JP, Ristanis SE, Soucacos PN. Botulinum A toxin for treatment of lower limb spasticity in cerebral palsy: Gait analysis in 49 patients. *Acta Orthop Scand*. 2003;74(6):749-755.

41. Polak F, Morton R, Ward C, Wallace WA, Doderlein L, Siebel A. Double-blind comparison study of two doses of botulinum toxin A injected into calf muscles in children with hemiplegic cerebral palsy. *Dev Med Child Neurol*. 2002;44(8):551-555.

42. Sutherland DH, Kaufman KR, Wyatt MP, Chambers HG. Injection of botulinum A toxin into the gastrocnemius muscle of patients with cerebral palsy: A 3-dimensional motion analysis study. *Gait Posture*. 1996;4(4):269-279.

43. Sutherland DH, Kaufman KR, Wyatt MP, Chambers HG, Mubarak SJ. Double-blind study of botulinum A toxin injections into the gastrocnemius muscle in patients with cerebral palsy. *Gait Posture*. 1999;10(1):1-9.

44. Svehlik M, Kraus T, Steinwender G, et al. Repeated Multilevel Botulinum Toxin A Treatment Maintains Long-Term Walking Ability in Children with Cerebral Palsy. *Ces A Slov Neurol A*

Neurochir. 2012;75(6):737-741.

45. Thompson NS, Baker RJ, Cosgrove AP, Corry IS, Graham HK. Musculoskeletal modelling in determining the effect of botulinum toxin on the hamstrings of patients with crouch gait. *Dev Med Child Neurol*. 1998;40(9):622-625.

46. Wissel J, Muller J, Baldauf A, et al. Gait analysis to assess the effects of botulinum toxin type A treatment in cerebral palsy: an open-label study in 10 children with equinus gait pattern. *Eur J Neurol*. 1999;6(4):S63-S67.

47. Zurcher AW, Molenaers G, Desloovere K, Fabry G. Kinematic and kinetic evaluation of the ankle after intramuscular injection of botulinum toxin A in children with cerebral palsy. *Acta Orthop Belg*. 2001;67(5):475-480.

Supporting Information

Table S1. Kinematic pelvic features extracted from 26 papers, identified through systematic literature search.

	26	27	28	29	30	31	32	18	33	34	35	5	36	37	38	39	9	40	41	42	43	44	45	46	6	47
Sagittal plane																										
Max. anterior tilt																							**x**			
Mean tilt							**v**	**x**	**v**								**x**					**v**				
ROM								**x**	**x**			**x**					**x**									
Coronal plane																										
Mean obliquity								**x**									**x**									
ROM								**x**	**x**								**x**									
Transverse plane																										
Mean rotation								**x**									**x**									
ROM								**v**	**x**								**x**									

V = Significantly changed post-BTX-A treatment; X = Not significantly changed post-BTX-A treatment; Max. = Maximal; ROM = Range Of Motion. Column headings represent numbers of the studies on BTX-A treatment in the reference list.

Table S2. Kinematic hip features extracted from 26 papers, identified through systematic literature search.

	26	27	28	29	30	31	32	18	33	34	35	5	36	37	38	39	9	40	41	42	43	44	45	46	6	47
Sagittal plane																										
Flexion angle at IC								**x**	**x**			**v**					**x**	**x**								
Max. ext. during St							**x**		**v**			**x**						**v**				**v**				
Flexion angle at TSt								**x**									**x**	**x**								
Max. flexion during Sw							**x**	**x**	**x**								**x**	**x**					**x**			
ROM during St								**x**									**x**									
ROM							**v**					**v**						**x**								
Coronal plane																										
Mean angle during St								**x**	**x**								**x**									
Mean angle during Sw								**x**	**v**								**x**									
Max. abd. angle																		**x**								
Max. add. angle																		**x**								
ROM								**x**									**x**									
Transverse plane																										
Angle at IC								**x**	**v**								**x**									
Angle at 50% of St								**x**									**x**									
Angle at TSt								**x**									**x**									
Angle at toe-off									**v**																	
Angle at 50% of Sw								**x**	**v**								**x**									
Mean rotation during St																						**v**				

V = Significantly changed post-BTX-A treatment; X = Not significantly changed post-BTX-A treatment; Max. = Maximal; ROM = Range Of Motion; ext. = extension; St = stance; Sw = swing; IC = initial contact; TSt = terminal stance; abd. = abduction; add. = adduction. Column headings represent numbers of the studies on BTX-A treatment in the reference list.

Table S3. Kinematic knee features extracted from 26 papers, identified through systematic literature search.

	26	27	28	29	30	31	32	18	33	34	35	5	36	37	38	39	9	40	41	42	43	44	45	46	6	47
Flex. angle at IC						**x**	**v**	**x**	**x**			**v**					**x**	**x**								**v**
Flex. angle during LR																		**x**								
Max. flex. during St								**x**	**x**								**x**			**x**						
Max. ext. during St		**v**				**x**	**v**	**v**	**x**			**x**		**x**				**v**		**x**		**x**	**v**		**x**	
Angle at toe-off								**x**																		
Flex. angle at TSt																	**v**	**x**								
Flex. angle at PSw																		**x**								
Max. flex. during Sw						**x**	**x**	**x**	**x**			**v**					**x**	**x**							**x**	
ROM during St								**v**									**x**									
ROM						**x**	**v**					**v**						**x**						**x**		
Timing of ext. motion								**v**									**x**									
Amount of delayed flex. in Sw								**v**									**x**									

V = Significantly changed post-BTX-A treatment; X = Not significantly changed post-BTX-A treatment; Max. = Maximal; ROM = Range Of Motion; Flex. = flexion; Ext. = extension; St = stance; Sw = swing; IC = initial contact; LR = loading response; TSt = terminal stance; PSw = pre-swing. Column headings represent numbers of the studies on BTX-A treatment in the reference list.

Table S4. Kinematic ankle features extracted from 26 papers, identified through systematic literature search.

	26	27	28	29	30	31	32	18	33	34	35	5	36	37	38	39	9	40	41	42	43	44	45	46	6	47
Angle at IC	**v**		**x**			**v**		**v**	**x**	**v**	**v**	**x**				**v**	**v**	**v**		**v**						**v**
Angles during 1R								**v**							**v**		**x**				**v**					
DF angle at 50% of St					**v**																					
DF ankle at TSt																		**v**								
Timing Max. DF in St								**v**									**v**								**v**	
Max. DF angle in St	**v**	**v**	**x**	**v**	**v**	**v**		**v**	**x**	**v**	**v**	**v**	**v**	**v**	**v**	**v**	**v**	**v**	**v**	**v**	**v**	**v**			**v**	**v**
Max. PF angle						**v**						**x**						**x**								
DF angle during PSw																		**v**								
DF angle at 50% of Sw					**v**			**v**		**v**							**v**									
Max. DF angle in Sw	**v**			**v**									**v**			**v**			**v**	**v**	**v**					
Max. PF angle in Sw			**x**								**x**				**v**											
ROM during Sw								**v**									**v**									
ROM during toe-off								**x**	**x**								**x**									
ROM						**x**						**v**				**v**		**v**						**x**		
Mean rotation in St								**v**									**x**									

V = Significantly changed post-BTX-A treatment; X = Not significantly changed post-BTX-A treatment; Max. = Maximal; ROM = Range Of Motion; IC = initial contact; 1R = first rocker; St = stance; Sw = swing; TSt = terminal stance; PSw = pre-swing; DF = dorsiflexion; PF = plantarflexion. Column headings represent numbers of the studies on BTX-A treatment in the reference list.

Table S5. Kinematic features of the foot in the transverse plane extracted from 26 papers, identified through systematic literature search.

	26	27	28	29	30	31	32	18	33	34	35	5	36	37	38	39	9	40	41	42	43	44	45	46	6	47
Mean progression angle during stance								**v**	**x**								**v**									
External rotation in swing																				**v**						

V = Significantly changed post-BTX-A treatment; X = Not significantly changed post-BTX-A treatment. Column headings represent numbers of the studies on BTX-A treatment in the reference list.

Table S6. Kinetic hip features extracted from 26 papers, identified through systematic literature search.

	26	27	28	29	30	31	32	18	33	34	35	5	36	37	38	39	9	40	41	42	43	44	45	46	6	47
Sagittal plane																										
Max. ext. moment in stance																	**x**									
Max. flex. moment in stance																	**x**									
Timing of 0Nm moment (% of GC)																	**x**									
Max. power absorption in stance																	**x**									
Max. power generation in stance																	**x**									
Max. power generation in first phase of stance												**v**														
Power generation at pre-swing																	**x**									
Coronal plane																										
Max. abd. moment in stance																	**x**									

V = Significantly changed post-BTX-A treatment; X = Not significantly changed post-BTX-A treatment; Max. = Maximal; ext. = extension; flex. = flexion; GC = gait cycle; abd. = abduction. Column headings represent numbers of the studies on BTX-A treatment in the reference list.

Table S7. Kinetic knee features extracted from 26 papers, identified through systematic literature search.

	26	27	28	29	30	31	32	18	33	34	35	5	36	37	38	39	9	40	41	42	43	44	45	46	6	47
Max. flexion moment in stance																	**X**									
Max. extension moment in stance																	**X**									
Max. power generation during stance												**V**					**X**									
Max. power absorption during (terminal) stance												**X**					**X**									

V = Significantly changed post-BTX-A treatment; X = Not significantly changed post-BTX-A treatment; Max. = Maximal. Column headings represent numbers of the studies on BTX-A treatment in the reference list.

Table S8. Kinetic ankle features extracted from 26 papers, identified through systematic literature search.

	26	27	28	29	30	31	32	18	33	34	35	5	36	37	38	39	9	40	41	42	43	44	45	46	6	47
1st max. PF moment (1st peak before MSt)			**x**							**x**							**x**									**v**
PF moment at the end of 1st double support phase																**v**										
PF moment at LR								**v**									**v**									
2nd max. PF moment (2nd peak before TO)			**x**					**v**		**x**							**v**									**v**
PF moment at begin of 2nd DS phase																**x**										
Max. PF moment in Mst - TSt								**v**																		
Max. moment																									**x**	
Timing of max. moment (% of GC)																									**v**	
M index											**v**															
M area											**v**															
AMQ					**v**											**x**										**v**
APQ					**v**											**x**										**v**
Max. power generation during GC	**x**									**x**		**x**				**x**						**v**			**x**	**x**
Max. generation in MSt																									**v**	**v**
Power generation at PSw								**x**									**x**									
Max. power absorption in early stance										**x**		**v**				**x**										
Power absorption at LR								**v**									**v**									
Generated work										**x**	**x**															
Absorbed work										**v**	**x**															

V = Significantly changed post-BTX-A treatment; X = Not significantly changed post-BTX-A treatment; Max. = Maximal; PF = plantarflexion; MSt = mid-stance; LR = loading response; TO = toe-off; MSt = midstance; TSt = terminal stance; GC = gait cycle; M index = equation 1 in reference [35]; M area = area beneath moment curve; AMQ = 1st moment peak / 2nd moment peak; APQ = 1st power generation peak / 2nd power generation peak; PSw = preswing. Column headings represent numbers of the studies on BTX-A treatment in the reference list.

Chapter 3

Identification of joint patterns during gait in children with cerebral palsy: a Delphi consensus study

Angela Nieuwenhuys
Sylvia Õunpuu
Anja Van Campenhout
Tim Theologis
Josse De Cat
Jean Stout
Guy Molenaers
Tinne De Laet
Kaat Desloovere

Developmental Medicine and Child Neurology 2016 Mar;58(3);306-13.

Abstract

Aim

This study aims to achieve an international expert consensus on joint patterns during gait for children with cerebral palsy (CP) by means of Delphi surveys.

Methods

In stage one, seven local experts drafted a preliminary proposal of kinematic patterns for each lower limb joint in the sagittal, coronal, and transverse plane. In stage two, thirteen experts from eight gait laboratories, four in the US and four in Europe, participated in a Delphi consensus study. Consensus was defined by a pre-set cut-off point of 75% agreement among participants.

Results

After the first stage, 44 joint patterns were presented in a first survey and 31 patterns reached consensus. Consensus improved to 47 out of 48 patterns in the third survey. Only one pattern, 'abnormal knee pattern during loading response', did not reach consensus. The expert panel agreed to define six patterns for the knee during swing, most of them representing characteristics of a stiff knee pattern.

Conclusion

The defined joint patterns can support clinical reasoning for children with CP as joint patterns during gait might be linked to different treatment approaches. Automating the classification process and incorporating additional trunk, foot, and electromyography features should be prioritized for the near future.

Introduction

Cerebral palsy (CP) is a heterogeneous condition not only in terms of etiology, but also with regards to clinical presentation[1]. Consequently, many forms of pathological gait can be identified and each requiring a specific treatment approach. Gait in CP is typically assessed through instrumented three dimensional gait analysis (3DGA), providing information on joint angles, moments and power. The outcome of 3DGA is highly valued with regards to treatment decision making[2,3]. A specific, detailed classification system for pathological gait using 3DGA data may thereby have many possible advantages. Apart from research applications, gait classifications can improve communication among healthcare workers by providing a tool for describing, evaluating and comparing gait between and within patients or groups of patients. Ultimately, it could aid lecturers teaching about gait in CP, serve as a tool for assessing treatment outcome, and potentially lead to a more in-depth understanding of the neurological cause of specific gait patterns, which may be associated with specific treatment indications.

Many gait classification systems have been developed by means of quantitative or qualitative construction techniques[4]. Still, several methodological and practical concerns remain. One of the main concerns regarding unsupervised quantitative construction techniques are the potential artificial groupings that may be produced[4]. Often, clinical interpretation of the patterns is complex, which constrains easy implementation in medical practice. Additionally qualitative classification systems, e.g. the classification for patients with hemiplegia by Winters et al.[5], have been criticized due to their lack of transparency regarding the construction process and limited agreement between clinicians[4,6]. Consensus approaches such as normative group theory or Delphi consensus studies could therefore be of major importance for studies that use a qualitative approach[7]. Content validity of many classification systems may also be compromised as deviations across the three anatomical planes are often not considered. Selecting an appropriate classification system may further be a challenge because they are frequently developed for a specific clinical subgroup, e.g. children with unilateral CP. All these concerns impede routine clinical application of many gait classification systems.

The final goal of the present study is to develop a new and automated gait classification system for all children with the spastic type of CP. As a first step, this paper will report the

results of a Delphi consensus study. The study aimed to achieve an international expert consensus on all clinically relevant joint patterns that should be incorporated in this classification system, ensuring transparency in the construction process and clinical interpretability of the defined joint patterns. In this first step, patterns are defined at the level of each joint separately, using objective 3DGA data.

Methodology

Ethical approval for the project was granted by the Medical Ethical Committee of University Hospitals Leuven, reference s56036. The development of the gait classification system comprised two stages. First, a local multidisciplinary team developed a preliminary proposal, identifying different kinematic and kinetic patterns for each lower limb joint across the three anatomical planes. Secondly, this proposal was reviewed by an international expert panel via three consecutive Delphi surveys. The target population for the proposed joint patterns includes ambulatory children with a predominantly spastic diagnosis of CP.

Stage 1 – preliminary proposal of classification system

Based on an extensive literature review and an available CP reference database, a local team of seven clinicians developed an initial list of joint patterns, which was published for the knee and ankle joint by Van Gestel et al.[8]. Patterns could comprise a single kinematic or kinetic deviation, either referring to a specific point in the gait cycle such as maxima, or referring to a deviation pointing to the overall shape or position of the waveform. No general instructions or guidelines were used as to what a pattern should look like or by how many features a pattern needed to be characterized. A detailed description of the expert team and the process that was undertaken to develop this initial proposal is available in Van Gestel et al.[8].

Stage 2 – Delphi consensus study

The Delphi technique is a consensus method that systematically structures the qualitative opinions of an expert panel by means of sequential surveys[9,10]. An international expert panel was consulted to provide their opinion on the relevance and content validity of all joint patterns that were defined in Stage 1. Only experts with more than ten years of experience in the field of CP and over five years of experience in teaching at international gait courses were eligible to participate in the study. Furthermore, experts had to undertake or interpret 3DGA data on a weekly basis. A range of different professions was ensured, especially when

multiple experts from the same gait laboratory were selected. Selected experts were invited in a group meeting to consider the proposed joint patterns from Stage 1. This meeting started by introducing the contents and objectives of the project, followed by a detailed presentation of all joint patterns, which were illustrated with examples using kinematic and kinetic data.

Afterwards, the first survey round was distributed and completed. The survey was followed by a group discussion of approximately two hours to clarify the main concerns and opinions. Following each survey, the expert panel was presented with the results of the previous round, which allowed each participant to consider their individual opinion in light of the panel's responses. It also permitted an anonymous and equal contribution of all experts without peer pressure. Furthermore, Delphi studies can function over long distances as frequent group meetings of all panel members were not feasible in practice.

Data Collection and Analysis

Surveys were created using web-based online software (www.kwiksurveys.com) and participant invitations were sent via email. During the second and third round, surveys were available online for seven weeks and one reminder was sent during that period. The first author anonymously analyzed and redrafted all surveys. Results of previous rounds were presented to the panel using descriptive statistics. With each question, participants were encouraged to add written comments and suggestions, which were provided in their original wording in the subsequent survey. Any form of suggestion was accepted, meaning that no instructions were given as to how a pattern should be shaped or by how many features a pattern could be characterized. Patterns were scored on a 5-point symmetric scale ranging from 'strongly agree' to 'strongly disagree'. In the third and final survey, participants were solely asked to 'agree' or 'disagree' about all patterns of a specific joint, providing a consensus on all patterns of this joint had already been reached during the previous rounds. A pre-set cut-off point of 75% agreement was defined as consensus, meaning 75% of participants rated 'agree' or 'strongly agree' for a pattern. In case a consensus was between 50% and 75% agreement, patterns were accommodated to the feedback and suggestions of participants and proposed for further discussion in the next survey. Patterns were removed when agreement was below 50%. Whenever three or more experts suggested a new pattern or an adaptation of an existing pattern, it was added to the classification system and proposed in the following survey. Patterns that were suggested only once or twice were questioned

separately. Figure 1 presents an overview of the Delphi process. Sample questions for each of three survey rounds are provided in the online supporting information.

Results

Stage 1 – preliminary proposal of classification system

The local team considered patterns for the pelvis and hip in the three anatomical planes, the knee and ankle in the sagittal plane, and the foot progression angle. In the end, 44 joint patterns were proposed. They are listed in Tables 1 and 2, and a full description of all patterns at this stage is provided in Table SI in the online supporting material. During this initial stage, no patterns were proposed for the hip in the coronal plane.

Stage 2 - Delphi consensus study

Results of the Delphi surveys are briefly summarized in Figure 1. Fourteen experts from eight gait laboratories, four in the US and four in Europe, met the inclusion criteria and were invited to participate. Thirteen experts accepted the invitation, among which pediatric orthopedic surgeons (n=5), kinesiologists and physiotherapists (n=4), and biomechanical scientists or engineers specialized in gait analysis or biomechanics in CP (n=4). In round two, one expert withdrew further participation due to a personal family emergency. Tables 1 and 2 describe the extent of agreement among experts for the joint patterns across all survey rounds. On four occasions during the first round, an expert indicated that a proposed joint pattern was not understood. Their responses for those specific patterns were considered as missing data.

Table 3 presents an overview of the principal changes that were made to all joint patterns. Most changes were made after the first survey and often terminology was adjusted to avoid confusion, e.g. specifying whether moments are internal or external moments. After round three, consensus was not reached on one knee pattern, namely ‘abnormal knee pattern during loading response’. A detailed definition of the final joint patterns after the last survey round is provided in table SII in the online supporting material.

During the group discussion, some important concerns were raised. First, different forms of error and uncertainty can arise when collecting 3DGA data. However, there was a unanimous consensus that 3DGA data could be reliably used given a thorough quality check, and given that data are collected by experienced professionals following a validated protocol.

Another point of discussion was the target patient population for the classification. Unless fore-foot, mid-foot, and hind-foot motion are adequately assessed with multi-segment foot models, misinterpretations can arise about the presented joint patterns, such as calcaneus gait. All experts agreed that multi-segment foot models, but also trunk models are not yet routinely used in all clinical motion analysis laboratories. Therefore, only data collected using well-known, conventional gait analysis protocols were considered to define joint patterns. This means that the currently presented classification requires children to have a rigid foot position in stance without severe foot deformities as well as sufficient foot clearance during swing. As a guideline, excessively increased or decreased joint angles as they are described for the different joint patterns generally refer to a deviation which is at least one standard deviation away from a reference database of typically developing children.

Six experts considered EMG during gait to be an important, valuable measure. Because quantifying and analyzing EMG data still holds many challenges, the experts unanimously agreed that EMG results should not yet be included in the classification system.

During the discussion, consensus could not be reached on the classification criteria of 'stiff knee'. Literature on the stiff knee also reflects heterogeneity in classification criteria and variable treatment outcomes of patients classified as stiff knee[11,12]. In round three, a unanimous preference was therefore given to define six easily recognizable patterns, including e.g. delayed peak knee flexion, which are in essence an enumeration of potential classification criteria of a stiff knee.

During the first survey, 31 additional joint patterns were suggested by only one or two experts and were thus not automatically added to the classification system. Their relevance was questioned in round two and only seven were deemed '(very) important' by at least 75% of the panel. Of these seven features, only 'reduced ankle power generation' was considered strictly essential to be incorporated in the classification system by at least 75% of participants. The seven features are described in Table SIII in the online material.

Table 1. Consensus on pathological joint patterns in sagittal plane improves across three Delphi rounds.

	Round 1	Round 2	Round 3
	Consensus (n)	Consensus (n)	Consensus (n)
Pelvis			
Normal pelvic motion/posture	12/13	9/9	9/9
Increased pelvic range ROM	8/11	8/9	9/9
Increased pelvic anterior tilt on average	11/13	9/9	9/9
Increased pelvic anterior tilt + increased ROM	9/11	8/9	9/9
Decreased pelvic anterior tilt (posterior tilt) on average	12/13	8/9	9/9
Decreased pelvic anterior tilt (posterior tilt) + increased ROM	8/11	7/9	9/9
Hip			
Normal hip motion	12/13	9/9	9/9
Increased hip power generation around toe-off (H3 burst)	4/13	na	na
Hip extension deficit	12/13	9/9	8/9
Hip extension deficit + H3 burst	4/13	na	na
Continuous excessive hip flexion	11/13	7/9	7/9
Continuous excessive hip flexion + H3 burst	5/13	na	na
Knee during stance phase			
Normal knee motion during stance	12/13	9/9	9/9
Abnormal pattern during loading response	na	na	5/9
Increased knee flexion at initial contact	12/13	9/9	na
Increased knee flexion at initial contact + earlier extension movement	9/13	4/9	na
Knee hyperextension	12/13	9/9	8/9
Knee hyperextension + increased knee flexion at initial contact	11/13	9/9	9/9
Increased flexion in midstance + internal flexion moment present	9/13	8/9	7/9
Increased flexion in midstance + internal extension moment present	10/13	8/9	8/9
Knee during swing phase			
Normal knee motion during swing	12/13	9/9	9/9
Stiff knee	11/13	na	na
Delayed peak knee flexion	na	5/5	9/9
Increased peak knee flexion	8/13	5/9	8/9
Increased + delayed peak knee flexion	na	3/5	8/9
Decreased peak knee flexion	na	5/5	9/9
Decreased + delayed peak knee flexion	na	5/5	9/9
Ankle during stance phase			
Normal ankle motion during stance	12/13	9/9	9/9
Horizontal second ankle rocker	11/13	7/9	9/9
Reversed second ankle rocker	11/12	8/9	9/9
Equinus	12/13	9/9	9/9
Calcaneus gait	9/13	9/9	9/9
Ankle during swing phase			
Normal ankle motion during swing	12/13	9/9	9/9
Insufficient prepositioning in terminal swing	11/12	9/9	9/9
Continuous plantarflexion during swing (drop foot)	11/13	9/9	9/9
Excessive dorsiflexion during swing	10/13	6/9	8/9

na = not applicable; ROM = range of motion; Terminology in the table is based on the terminology that was used in the final Delphi survey. Light grey areas indicate consensus (75% agreement); dark grey areas indicate 100% agreement.

Table 2. Consensus on pathological joint patterns in coronal and transverse plane improves across three Delphi rounds.

	Round 1	Round 2	Round 3
	Consensus (n)	Consensus (n)	Consensus (n)
Pelvis in coronal plane			
Normal pelvic motion/posture	11/13	9/9	9/9
Pelvic instability	8/12	na	na
Increased pelvic ROM	na	8/9	9/9
Continuous pelvic elevation (up)	na	8/9	8/9
Continuous pelvic elevation (down)	na	8/9	8/9
Hip in coronal plane			
Normal hip motion	11/13	9/9	9/9
Pathological motion	7/12	na	na
Excessive hip abduction in swing	na	8/9	8/9
Continuous excessive hip abduction	na	6/9	7/9
Continuous excessive hip adduction	na	7/9	7/9
Pelvis in transverse plane			
Normal pelvic motion/posture	12/13	9/9	9/9
Increased pelvic ROM	8/13	7/9	9/9
Excessive pelvic external rotation during the gait cycle	na	9/9	9/9
Excessive pelvic internal rotation during the gait cycle	na	9/9	9/9
Hip in transverse plane			
Normal hip motion	11/13	9/9	9/9
Hip deviation explained through pelvic deviation	6/13	na	na
Excessive hip external rotation during the gait cycle	11/13	9/9	9/9
Excessive hip internal rotation during the gait cycle	11/13	9/9	9/9
FPA			
Normal FPA	12/13	9/9	8/9
Outtoeing	10/13	9/9	8/9
Intoeing	11/13	9/9	8/9

na = not applicable; ROM = range of motion; FPA = foot progression angle; Terminology in the table is based on the terminology that was used in the final Delphi survey. Light grey areas indicate consensus (75% agreement); dark grey areas indicate 100% agreement.

Table 3. Principal adaptations to the classification system across three Delphi rounds.

	Round 1		Round 2		Round 3	
	Consensus /total number of patterns	***Qualitative analysis + discussion meeting***	***Consensus /total number of patterns***	***Qualitative analysis***	***Consensus /total number of patterns***	***Qualitative analysis***
PELVIS						
Sagittal plane	4/6	- Delete the feature 'single or double bump pattern throughout stance' - Replace 'pelvic instability' by 'increased range of motion'	6/6	- no changes	6/6	- no changes
Coronal plane	1/2	Two patterns were added: - Pelvic elevation - Pelvic depression - Replace 'pelvic instability' by 'increased range of motion'	4/4	- no changes	4/4	- Necessary to evaluate pelvic deviations specific to stance and swing?
Transverse plane	1/2	Two patterns were added: - Excessive external rotation - Excessive internal rotation	4/4	- no changes	4/4	- Necessary to evaluate pelvic deviations specific to stance and swing?
HIP						
Sagittal plane	3/6	- Delete all patterns with the feature 'increased hip power generation around toe-off (=H3 burst)'.	3/3	- no changes	3/3	- no changes
Coronal plane	1/2	Three patterns were added: - Excessive hip abduction in swing - Continuous excessive hip abduction - Continuous excessive hip adduction	3/4	- 'Continuous excessive hip abduction' did not reach consensus, but was proposed again in round 3 as no expert disagreed and no written remarks were made	4/4	- Necessary to evaluate additional hip deviations specific to stance and swing?
Transverse plane	3/4	- Delete 'hip deviation explained through pelvic deviation' as this is not a pure hip pattern	3/3	- no changes	3/3	- no changes

Table 3. Continued.

KNEE						
Sagittal during stance	5/7	- Rephrase the patterns 'Increased knee flexion + knee flexion moment present/absent' to make the meaning more clear; change to 'internal knee flexion moment present' or 'internal knee extension moment present' - Suggestion to rephrase the pattern 'increased knee flexion at initial contact and earlier knee extension movement'	6/7	- No consensus for 'Increased knee flexion at initial contact and earlier knee extension movement'. However, no expert disagreed and the proposed terminology was preferred over the alternative suggestion 'increased knee flexion at initial contact and early minimum knee flexion'. A merge of this pattern with 'increased knee flexion at initial contact' is proposed.	5/6	- The merged pattern 'Abnormal pattern during loading response' did not reach consensus.
Sagittal during swing	2/3	- During the group discussion, it was decided to ask the expert panel about possible alternatives for a classification of stiff knee. Either the current configuration was maintained or a classification defining six patterns during swing would be defined.	---	- Five out of nine experts preferred to define six patterns in swing. These patterns were suggested in round 3. It was argued that the four other participants, who initially preferred the stiff knee pattern, also agreed that in any case 'decreased peak knee flexion' and 'delayed timing of peak knee flexion' should be included in the classification system.	6/6	- All experts agreed to continue with six patterns and consensus was reached on all of them.
ANKLE						
Sagittal during stance	4/5	- Add the features 'increased slope towards dorsiflexion' to the pattern	5/5	- no changes	5/5	- no changes
Sagittal during swing	4/4	- no changes	3/4	- 'excessive dorsiflexion in swing' was proposed again, as no expert disagreed and no textual remarks or suggestions to change the pattern were made.	4/4	- no changes
FPA	3/3	- no changes	3/3	- no changes	3/3	- no changes

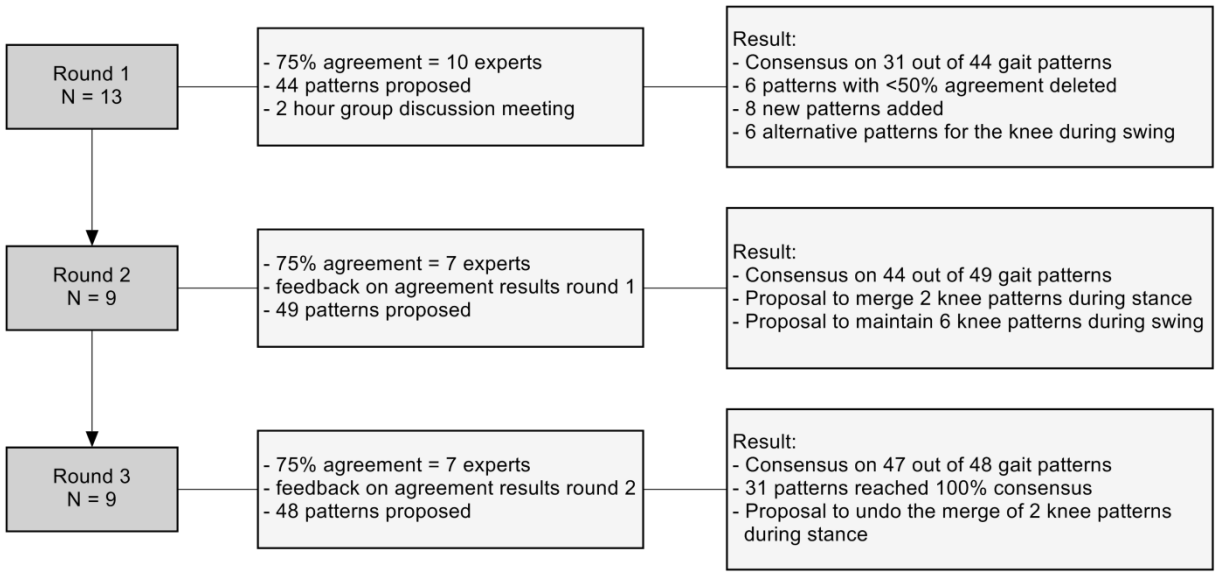


Figure 1. Flow chart summarizes expert participation and different steps of the Delphi process.

Discussion

In this paper, we documented a Delphi consensus study, which introduced joint patterns for the lower limb joints across the three anatomical planes, based on objective 3DGA data. An international expert consensus was achieved on all patterns, except for one knee pattern during stance. Patterns rely mainly on kinematics, which make them also relevant for children using e.g. Kaye Walkers, for whom kinetics are unavailable.

Consensus was reached on all patterns for the pelvis and hip across the three anatomical planes. Wren et al.[13] evaluated the prevalence of excessive hip flexion, adduction, and internal hip rotation. They evaluated hip adduction and internal rotation during stance whereas the current study evaluates these patterns across the entire gait cycle. Although it was proposed by some experts to define pelvis and hip patterns specific to the stance phase, no group consensus was reached to include them.

In the sagittal plane seven patterns were defined for the knee during stance and consensus was reached on all but one. It was argued that in round two ‘increased knee flexion at initial contact’ reached 100% consensus and no expert disagreed on ‘increased knee flexion at initial contact and earlier knee extension movement’. It is therefore proposed to continue the final classification with those two original patterns. In the past, Simon et al.[14] defined two genu recurvatum patterns. These two patterns were differentiated among others by the

interdependency of the motion of the knee with ankle and trunk movements. As such, this distinction could not have been detected in the current study where patterns were considered for each joint separately. Sutherland et al.[15] have described the most commonly used knee patterns. Sutherland's jump knee, recurvatum knee, and crouch knee clearly show similarities to the knee patterns of the current study. This is not surprising since joint patterns in this study are, apart from expert opinions, also originally founded on an extensive literature review. It also explains why patterns can be characterized by findings from different studies. For instance, the two knee hyperextension patterns of the current study include similarities with Sutherland's recurvatum knee, but they also include kinetic features, which have been identified by Lin et al.[16].

A stiff knee pattern during swing was consciously not adopted in our study. Instead, six easily recognizable patterns, the majority of which are potential classification criteria of a stiff knee, were defined and agreed upon by the panel. Consequently, communication about knee patterns during swing could hopefully be facilitated.

Consensus was easily reached on all ankle patterns, both during stance and swing. Equinus and calcaneus gait were also reported by Wren et al.[13]. The current study adopts a more narrow definition of equinus and as such, lower prevalence numbers could be expected when applying the currently presented definition of equinus. Schmidt-Rohlfing et al.[17] defined three ankle patterns based on EMG evaluations of m. tibialis anterior, m. gastrocnemius and m. soleus. It might be interesting to explore how their work could be related to kinematic joint patterns.

A frequently highlighted limitation for Delphi surveys is the subjective nature of the technique[9,18]. However, a Delphi approach was important as it ensured the necessary clinical interpretability of the defined joint patterns, in contrast to many quantitative construction techniques such as k-means cluster analysis, where clinical interpretation is often limited. It further guaranteed transparency in the development process of the joint patterns. Another common limitation of Delphi studies relates to the size and subjective selection process of the expert panel[18,9]. Given the specialized nature of the topic of the study, it was appropriate to recruit a relatively small expert panel. One out of thirteen experts was unfortunately not able to participate any further. With nine experts participating in the following rounds, response rate was stable and unexplained drop-out remained under 30%, as is suggested by Sumsion et al.[19]. Furthermore, a multi-center and multidisciplinary panel was

recruited, with experts who already have longstanding experience concerning gait in CP. In addition, patterns in this study are supported by literature findings. As a result, we are confident that the results of this Delphi survey will be valuable both in clinical and research settings.

This study implemented a modified Delphi approach as the first survey started from a preliminary proposal of joint patterns that were defined by a local expert team[8]. Over 40 new patterns were suggested after the first survey round, yet only twelve patterns were agreed upon after the final survey. In the deletion or addition of new patterns, experts were challenged to find a balance between including sufficient clinical detail and ensuring practical feasibility of the final classification system. The difficulty in finding this balance could be reflected for instance in the disagreement between the experts to define additional patterns for the pelvis and hip. For example, it was suggested in the final round to define pelvis and hip patterns specific to stance and swing phase in the coronal and transverse plane. However, 'pelvic hiking in swing' and 'excessive hip adduction in swing' were, among others, all presented to the panel and considered by the majority of experts to be not strictly essential in the classification system. In part, this might be explained by the compensatory nature of many coronal and transverse plane patterns. Perhaps they were considered to be not strictly essential in the classification system because they can often not be directly impacted by treatment. Even though the final number of patterns is still quite large and might seem complex, they are considered relevant as they might be linked to different treatment approaches. A possible solution is to reduce the number of joint patterns by focusing only on those patterns that can be directly affected by treatment. In further research, a large group of CP patients will be classified to examine the prevalence of each pattern and the likelihood of this prevalence being influenced by specific treatment approaches. We might then explore the possibility to identify two categories of patterns: primary joint patterns which reflect problems that can be directly impacted by treatment, and secondary joint patterns, which are compensatory and are expected to be normalized after treating the primary deviation. Because there is still much debate about the cause and effect relationship between different joints, especially in the coronal and transverse plane, these categories cannot be proposed yet.

Another solution to handle a large number of patterns is to automate this classification process by means of supervised learning techniques. To this end, Bayesian networks have been introduced in CP research as a means to combine a quantitative classification technique with

qualitative, clinical information[8,20]. Di Lello et al. have recently demonstrated promising classification results using such a Bayesian approach (article submitted in Transactions on Neural Systems & Rehabilitation Engineering). This automated approach will be able to provide a 'pattern profile' for each patient, which could be linked with clinical measures of muscle weakness and spasticity to improve our understanding about the etiology of the joint patterns.

In conclusion, this is the first study reporting a Delphi approach to define joint patterns for children with CP. Still, the patterns for this classification system should not be considered final at this stage. The classification system does not yet incorporate trunk motion, multi-segment foot evaluations, nor does it include EMG results. Defining additional trunk, foot, and EMG features could further characterize the currently presented joint patterns and should be prioritized for the near future.

Acknowledgements

The authors would like to thank J.R. Davids, M. Schwartz, R. Davis, and all others who have participated confidentially in the different surveys. Our gratitude also goes to L. Van Gestel for her dedicated work during the first stage of this study and her help during the start-up of the consensus study. A. Nieuwenhuys is supported by an OT project of KU Leuven University (OT/12/100).

References

1. Rosenbaum P, Paneth N, Leviton A, et al. A report: the definition and classification of cerebral palsy April 2006. *Dev Med Child Neurol Suppl*. 2007;109(April):8-14.
2. Lofterød B, Terjesen T. Results of treatment when orthopaedic surgeons follow gait-analysis recommendations in children with CP. *Dev Med Child Neurol*. 2008;50(7):503-509.
3. Wren TAL, Otsuka NY, Bowen RE, et al. Influence of gait analysis on decision-making for lower extremity orthopaedic surgery: Baseline data from a randomized controlled trial. *Gait Posture*. 2011;34(3):364-369.
4. Dobson F, Morris ME, Baker R, Graham HK. Gait classification in children with cerebral palsy: A systematic review. *Gait Posture*. 2007;25(1):140-152.
5. Winters TF, Gage JR, Hicks R. Gait patterns in spastic hemiplegia in children and young adults. *J Bone Joint Surg Am*. 1987;69:437-441.
6. Dobson F, Morris M, Baker R, Wolfe R, Graham H. Clinician agreement on gait pattern ratings in children with spastic hemiplegia. *Dev Med Child Neurol*. 2006;48(6):429-435.
7. Jones J, Hunter D. Qualitative Research: Consensus methods for medical and health services

research. *BMJ*. 1995;311:376-380.

8. Van Gestel L, De Laet T, Di Lello E, et al. Probabilistic gait classification in children with cerebral palsy: A Bayesian approach. *Res Dev Disabil*. 2011;32(6):2542-2552.
9. Hasson F, Keeney S, McKenna H. Research guidelines for the Delphi survey technique. *J Adv Nurs*. 2000;32(4):1008-1015.
10. Boulkedid R, Abdoul H, Loustau M, Sibony O, Alberti C. Using and reporting the Delphi method for selecting healthcare quality indicators: A systematic review. *PLoS One*. 2011;6(6).
11. Goldberg SR, Ounpuu S, Arnold AS, et al. Kinematic and kinetic factors that correlate with improved knee flexion following treatment for stiff-knee gait. *J Biomech*. 2006;39(4):689-698.
12. Dreher T, Wolf SI, Hagmann S, et al. Long-term results after distal rectus femoris transfer as a part of multilevel surgery for the correction of stiff-knee gait in spastic diplegic cerebral palsy. *J Bone Joint Surg Am*. 2012;142(e(142)):1-10.
13. Wren TAL, Rethlefsen S, Kay RM. Prevalence of specific gait abnormalities in children with cerebral palsy: influence of cerebral palsy subtype, age, and previous surgery. *J Pediatr Orthop*. 2005;25(1):79-83.
14. Simon S. Genu recurvatum in spastic cerebral palsy. Report on findings by gait analysis. *J bone Jt Surg - Am Vol*. 1978;60(7):882-894.
15. Sutherland DH, Davids JR. Common Gait Abnormalities of the Knee in Cerebral Palsy. *Clin Orthop Relat Res*. 1993;288:139-147.
16. Lin C, Guo L, Su F, Chou Y, Cherng R. Common abnormal kinetic patterns of the knee in gait in spastic diplegia of cerebral palsy. *Gait Posture*. 2000;11:224-232.
17. Schmidt-Rohlfing B, Bergamo F, Williams S, et al. Interpretation of surface EMGs in children with cerebral palsy: An initial study using a fuzzy expert system. *J Orthop Res*. 2006;24(3):438-447.
18. Keeney S, Hasson F, McKenna H. Consulting the oracle: Ten lessons from using the Delphi technique in nursing research. *J Adv Nurs*. 2006;53(2):205-212.
19. Sumsion T. The Delphi Technique: An adaptive research tool. *Br J Occup Ther*. 1998;61(4):153-156.
20. Zhang B, Zhang Y, Begg RK. Gait classification in children with cerebral palsy by Bayesian approach. *PATTERN Recognit*. 2009;42(4, SI):581-586.

Supporting Information

S1 Appendix: Sample question for each Delphi round

ROUND 1

Example 1

Ankle - sagittal plane

Statements below propose classes for classification of the ankle in the sagittal plane, in stance and in swing.

Do you feel there are additional clinically relevant (sub)classes that are not stated? Or perhaps you feel there are some redundant classes?
Please indicate your concerns in the text boxes below each question if you have any comments, additional remarks or other ideas regarding the suggested classes.

For the motion of the ***ankle*** in the sagittal plane ***in stance***, we can distinguish 5 separate classes:

0 - Normal ankle in stance
1 - Horizontal second ankle rocker
2 - Reversed second ankle rocker
3 - Equinus
4 - Calcaneus gait

Can you please indicate to what extent you agree with each class as being an important, valid class for the motion of the ***ankle in stance***.

	Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree
0 - Normal ankle in stance	○	○	○	○	○
1 - Horizontal second ankle rocker	○	○	○	○	○
2 - Reversed second ankle rocker	○	○	○	○	○
3 - Equinus	○	○	○	○	○
4 - Calcaneus gait	○	○	○	○	○

General comments, remarks or ideas

Example 2

According to your knowledge, how important and clinically relevant is it to incorporate electromyography (EMG) results from lower limb muscles (collected during gait analysis) into a gait classification system?

- Very important
- Important
- Neutral
- Unimportant
- Very unimportant

ROUND 2

Sagittal plane - Ankle in stance

"For the motion of the ***ankle in stance in the sagittal plane***, we can distinguish 5 separate classes. Please indicate to what extent you agree with each class as being an important, valid class for the motion of the ankle in stance."

	strongly agree	agree	neutral	disagree	strongly disagree	
Normal ankle in stance	**62%**	**31%**	8%	0%	0%	
Horizontal second ankle rocker	**23%**	**62%**	15%	0%	0%	
Reversed second ankle rocker	**42%**	**50%**	8%	0%	0%	***(n = 12)***
Equinus	**54%**	**38%**	8%	0%	0%	
Calcaneus gait	***38%***	***31%***	23%	8%	0%	

Suggestions:

- Presence / absence of the first rocker **(n = 2)**
- Reduced range of motion during stance
- Reduced plantar flexion during third rocker **(n = 2)**
- Consider ankle angle at initial contact
- Reduced ankle power generation
- A double bump pattern in stance

Other remarks:

- *The presence of foot deformity may influence the classification of reversed second ankle rocker, equinus and calcaneus gait, as mid-foot motion is presently not adequately assessed.* ***(n = 4)***

- *Each pathological pattern can be extended with a sub-pattern referring to the third ankle rocker.*

In light of the results presented above, please indicate to what extent you agree with each pattern as being an important, clinically relevant pattern for the motion of the *ankle in stance in the sagittal plane*.

	Strongly agree	Agree	Neutral	Disagree	Strongly disagree
Normal ankle in stance	○	○	○	○	○
Horizontal second ankle rocker	○	○	○	○	○
Reversed second ankle rocker	○	○	○	○	○
Equinus: - continuous plantarflexion throughout stance	○	○	○	○	○
Calcaneus gait: - increased slope towards dorsiflexion during stance OR - a peak > 20° of dorsiflexion	○	○	○	○	○

ROUND 3

Sagittal plane - Ankle in stance

"Please indicate to what extent you agree with each class as being an important, valid class for the motion of the ***ankle in stance in the sagittal plane***."

	strongly agree	agree	neutral	disagree	strongly disagree
Normal ankle in stance	**88.89%**	**11.11%**	0.00%	0.00%	0.00%
Horizontal second ankle rocker	**55.56%**	**22.22%**	22.22%	0.00%	0.00%
Reversed second ankle rocker	**66.67%**	**22.22%**	11.11%	0.00%	0.00%
Equinus	**88.89%**	**11.11%**	0.00%	0.00%	0.00%
Calcaneus gait	**77.78%**	**22.22%**	0.00%	0.00%	0.00%

For your convenience, a detailed overview of the patterns of the ankle in stance in the sagittal plane is presented once more below.

Do you agree with all these patterns and their classification rules?

- ○ I agree
- ○ I disagree

ANKLE stance – Sagittal plane

Normal ankle in stance
- No or minor gait deviations

Horizontal second ankle rocker
- Horizontal pattern of second ankle rocker from loading response (10%) to start push-off (slope < 5°)

Reversed second ankle rocker
- Descending or reversed pattern of second ankle rocker from loading response (10%) to start push-off (slope >= 5°).

Equinus
- Continuous plantarflexion (x<0°) throughout stance

Calcaneus gait
- Increased slope towards dorsiflexion during stance
  OR
- A peak >= 20° of dorsiflexion

Next Page

Option: I agree

Sagittal plane - Ankle in stance

If you have any further remarks or suggestions regarding content, terminology, ... for the patterns or their classification rules, please indicate them below.

ANKLE stance – Sagittal plane

Normal ankle in stance
- No or minor gait deviations

Horizontal second ankle rocker
- Horizontal pattern of second ankle rocker from loading response (10%) to start push-off (slope < 5°)

Reversed second ankle rocker
- Descending or reversed pattern of second ankle rocker from loading response (10%) to start push-off (slope >= 5°).

Equinus
- Continuous plantarflexion (x<0°) throughout stance

Calcaneus gait
- Increased slope towards dorsiflexion during stance
  OR
- A peak >= 20° of dorsiflexion

Option: I disagree

Can you please specify? --- Sagittal plane - ankle in stance

Which pattern or patterns do you disagree with?

- [ ] Normal ankle in stance
- [ ] Horizontal second ankle rocker
- [ ] Reversed second ankle rocker
- [ ] Equinus
- [ ] Calcaneus gait

Please state the reason(s) for your disagreement?

If you have any further remarks or suggestions regarding content, terminology, ..., please indicate them below.

Table S1. Stage 1: Preliminary proposal of classification system (before first Delphi survey).

		Joint pattern	Full description and criteria
Sagittal plane	**Pelvis**	Normal pelvic posture/motion	No or minor gait deviations observable in the pelvis in the sagittal plane
		Pelvic instability	Single or double bump pattern of pelvis motion throughout stance OR/AND increased range of motion throughout the gait cycle
		Increased pelvic anterior tilt	Increased pelvic anterior tilt on average
		Increased pelvic anterior tilt + Pelvic instability	Increased pelvic anterior tilt on average AND single or double bump pattern of pelvis motion throughout stance and/or increased range of motion throughout the gait cycle
		Decreased pelvic tilt	Decreased pelvic tilt on average
		Decreased pelvic tilt + Pelvic instability	Decreased pelvic tilt on average AND single or double bump pattern of pelvis motion throughout stance and/or increased range of motion throughout the gait cycle
	Hip	Normal hip motion	No or minor gait deviations observable in the hip in the sagittal plane
		H3 burst	Increased hip power generation around toe-off
		Hip extension deficit	At least two of the following characteristics: (1) decreased hip extension in stance (hip extension deficit), (2) decreased hip range of motion in stance, (3) delayed timing of zero hip moment or decreased hip flexion moment
		Hip extension deficit + H3 burst	H3 burst AND at least two of the following characteristics: (1) decreased hip extension in stance (hip extension deficit), (2) decreased hip range of motion in stance, (3) delayed timing of zero hip moment or decreased hip flexion moment
		Continuous excessive hip flexion	Continuous excessive hip flexion angle (above normal curve) throughout at least 90% of the gait cycle AND hip flexion angle continuously above 0°
		Continuous excessive hip flexion + H3 burst	H3 burst AND continuous excessive hip flexion angle (above normal curve) throughout at least 90% of the gait cycle AND hip flexion angle continuously above 0°
	Knee	Normal knee in stance	No or minor gait deviations observable in the knee in stance in the sagittal plane
	(during stance)	Increased knee flexion at initial contact	
		Increased knee flexion at initial contact + earlier knee extension movement	
		Knee hyperextension	At least two of the following characteristics: (1) increased knee extension (to knee hyperextension) in mid- or late stance, (2) earlier knee extension movement in stance, (3) increased (/excessive) knee flexion moment

Table S1. Continued.

Knee (during stance)	Knee hyperextension + increased knee flexion at initial contact	Increased knee flexion at initial contact AND at least two of the following characteristics: (1) increased knee extension (to knee hyperextension) in mid- or late stance, (2) earlier knee extension movement in stance, (3) increased (/excessive) knee flexion moment
	Increased knee flexion + knee flexion moment present	Increased knee flexion in midstance (= no normal knee angle in extension of slight flexion in midstance) AND at least one of the following characteristics: (1) increased knee flexion at initial contact, (2) some knee flexion moment preserved (= present for at least 1/3rd of stance phase), (3) increased peak knee extension moment during loading response
	Increased knee flexion + knee flexion moment absent	Increased knee flexion in midstance (= no normal knee angle in extension of slight flexion in midstance) AND at least one of the following characteristics: (1) increased knee flexion at initial contact, (2) knee flexion moment absent or present for less than 1/3rd of stance phase, (3) increased peak knee extension moment during loading response
Knee	Normal knee in swing	No or minor gait deviations observable in the knee in swing in the sagittal plane
(during swing)	Stiff knee	At least three of the following characteristics: (1) decreased knee range of motion between maximal knee extension in stance and peak knee flexion in swing, (2) delayed timing of peak knee flexion in swing, (3) decreased peak knee flexion in swing, (4) decreased knee flexion velocity around toe-off, (5) increased knee extension moment in double support (10% before toe-off to toe-off)
	Increased peak knee flexion in swing	
Ankle	Normal ankle in stance	No or minor gait deviations observable in the ankle in stance in the sagittal plane
(during stance)	Horizontal second ankle rocker	Cessation of tibial forward progression in second rocker (from loading response (10%) to start push-off) leading to a flat or horizontal pattern of second rocker (=horizontal curve ± 5°)
	Reversed second ankle rocker	Cessation of tibial forward progression in second rocker (from loading response (10%° to start push-off) followed by reversal of tibia movement leading to a reversed or "descending" pattern of second rocker
	Equinus	Continuous plantarflexion (x<0°) throughout stance
	Calcaneus gait	Continuous increased dorsiflexion throughout stance (with at least a peak ≥ 20°)
Ankle	Normal ankle in swing	No or minor gait deviations observable in the ankle in swing in the sagittal plane
(during swing)	Insufficient prepositioning in terminal swing	Ankle plantarflexion at initial contact at the end of the gait cycle (greater than normal values!)
	Continuous plantarflexion in swing	Prolonged plantarflexion for most of the swing phase (at least hindering foot clearance around 90% of the gait cycle) AND ankle plantarflexion at initial contact at the end of the gait cycle greater than normal values.
	Excessive dorsiflexion in swing	Increased dorsiflexion in swing for at least 1/3rd of the swing phase

Table S1. Continued.

Coronal plane	**Pelvis**	Normal pelvic posture/motion	No or minor gait deviations observable in the pelvis in the coronal plane
		Pelvic instability	Increased pelvic range of motion in the coronal plane
	Hip	Normal hip motion	No or minor gait deviations observable in the hip in the coronal plane
		Pathological motion	Clinically relevant deviation in motion or posture of the hip
Transverse plane	**Pelvis**	Normal pelvic posture/motion	No or minor gait deviations observable in the pelvis in the transverse plane
		Increased pelvic motion amplitude	Increased pelvic range of motion in the transverse plane
	Hip	Normal hip motion	No or minor gait deviations observable in the hip in the transverse plane
		Hip posture/motion explained through pelvic posture/motion	
		Hip exorotation	Excessive hip external rotation on average during entire gait cycle
		Hip endorotation	Excessive hip internal rotation on average during entire gait cycle
	Foot	Normal foot progression angle	No or minor gait deviations observable in the foot in the transverse plane
		Outtoeing	Excessive external foot progression on average during stance
		Intoeing	Excessive internal foot progression on average during stance

Table S2. Stage 2: Final overview of joint patterns and their criteria after last Delphi survey.

		Joint pattern	Full description and criteria
Sagittal plane	**Pelvis**	Normal pelvic posture/motion	
		Increased range of motion	
		Increased pelvic anterior tilt on average	
		Increased pelvic anterior tilt + increased range of motion	
		Decreased pelvic tilt (posterior tilt) on average	
		Decreased pelvic tilt (posterior tilt) + increased range of motion	
	Hip	Normal hip motion	
		Hip extension deficit	At least two of the following characteristics: (1) decreased hip extension in stance, (2) decreased hip range of motion in stance, (3), delayed timing of zero hip moment or decreased hip flexion moment
		Continuous excessive hip flexion	Excessive hip flexion throughout at least 90% of the gait cycle AND hip flexion angle continuously above 0°
	Knee	Normal knee in stance	
	(during stance)	Increased knee flexion at initial contact	
		Increased knee flexion at initial contact + earlier knee extension movement	
		Knee hyperextension	At least two of the following characteristics: (1) increased knee extension (to knee hyperextension) in mid- or late stance, (2) earlier knee extension movement in stance, (3) excessive knee flexion moment in mid- or late stance
		Knee hyperextension + increased knee flexion at initial contact	Increased knee flexion at initial contact AND at least two of the following characteristics: (1) increased knee extension (to knee hyperextension) in mid- or late stance, (2) earlier knee extension movement in stance, (3) excessive knee flexion moment in mid- or late stance
		Increased knee flexion in midstance + internal knee flexion moment present	Increased knee flexion in midstance: no normal knee angle in extension in midstance AND internal knee flexion moment is present for at least 1/3rd of stance phase
		Increased knee flexion in midstance + internal knee extension moment present	Increased knee flexion in midstance: no normal knee angle in extension in midstance AND internal knee extension moment is present for at least 2/3rd of stance phase

Table S2. Continued.

	Knee	Normal knee in swing	
	(during swing)	Delayed peak knee flexion	
		Increased peak knee flexion	
		Increased + delayed peak knee flexion	
		Decreased peak knee flexion	
		Decreased + delayed peak knee flexion	
	Ankle	Normal ankle in stance	
	(during stance)	Horizontal second ankle rocker	Horizontal pattern of second ankle rocker from loading response (10%) to start push-off (slope <5°)
		Reversed second ankle rocker	Descending pattern of second ankle rocker from loading response (10%) to start push-off (slope ≥-5°)
		Equinus	Continuous plantarflexion (x<0°) throughout stance
		Calcaneus gait	Increased slope towards dorsiflexion during stance OR a peak ≥ 20° of dorsiflexion
	Ankle	Normal ankle in swing	
	(during swing)	Insufficient prepositioning in terminal swing	Ankle plantarflexion at initial contact at the end of the gait cycle, which is greater than normal values
		Continuous plantarflexion in swing (drop foot)	Excessive plantarflexion for most of the swing phase, at least hindering foot clearance around 90% of the gait cycle AND ankle plantarflexion at initial contact at the end of the gait cycle, which is greater than normal values
		Excessive dorsiflexion in swing	Increased dorsiflexion in swing for at least 1/3rd of the swing phase
Coronal plane	**Pelvis**	Normal pelvic posture/motion	
		Increased pelvic range of motion	
		Continuous pelvic elevation (up)	
		Continuous pelvic depression (down)	
	Hip	Normal hip motion	
		Excessive hip abduction in swing	
		Continuous excessive hip abduction	
		Continuous excessive hip adduction	

Table S2. Continued.

Transverse plane	**Pelvis**	Normal pelvic posture/motion	
		Increased pelvic range of motion	
		Excessive pelvic external rotation during the gait cycle	
		Excessive pelvic internal rotation during the gait cycle	
	Hip	Normal hip motion	
		Excessive hip external rotation during the gait cycle	
		Excessive hip internal rotation during the gait cycle	
	Foot	Normal foot progression angle	
		Outtoeing	Excessive external foot progression on average during stance
		Intoeing	Excessive internal foot progression on average during stance

Table S3. Expert opinion on the importance of additional gait patterns and features in final Delphi round.

	Option A 'absolutely essential'	Option B 'not strictly essential'	Option C 'not important'
PELVIS in sagittal plane			
Distinction between single/double bump	6/9	3/9	0/9
HIP in sagittal plane			
Decreased hip power generation around toe-off	5/9	4/9	0/9
ANKLE during stance in sagittal plane			
Presence/absence first ankle rocker	6/9	3/9	0/9
Reduced ankle power generation	7/9	2/9	0/9
PELVIS in coronal plane			
Pelvic elevation/depression during stance	3/9	6/9	0/9
Pelvic hiking in swing	2/9	7/9	0/9
HIP in coronal plane			
Excessive adduction in swing	2/9	6/9	1/9

Option A: I agree, this feature is important. I consider this feature absolutely essential to be incorporated as an additional pattern in the gait classification system.
Option B: I agree, this feature is important. It should be specified whenever possible to further clarify the pattern, but it is not strictly essential to incorporate it as an additional pattern in the classification system.
Option C: I disagree, this feature is not so important that it should be incorporated in the gait classification system;
Light grey area indicates consensus (75% agreement).

Chapter 4

Inter- and intrarater clinician agreement on joint patterns during gait in children with cerebral palsy

Angela Nieuwenhuys
Eirini Papageorgiou
Guy Molenaers
Davide Monari
Tinne De Laet
Kaat Desloovere

Developmental Medicine and Child Neurology [in revision]

Abstract

Aim

This study aimed to quantify the inter- and intrarater agreement with which clinicians can recognize joint patterns during gait in children with spastic cerebral palsy (CP), which were recently specified by a Delphi consensus study. It also examined whether experience with three-dimensional gait analysis (3DGA) is a prerequisite for using the patterns.

Methods

The experimental group consisted of 82 CP patients (57 boys; uni-/bilateral CP (n=27/55); GMFCS I-III; mean age 9 years, 5 months). Thirty-two clinical raters were split into an 'experienced' and 'inexperienced' rater group. In two classification rounds, each rater was asked to classify 3DGA reports of 27 or 28 patients. Inter- and intrarater agreement on the 49 joint patterns was estimated using percentage of agreement and kappa statistics.

Results

Twenty-eight raters completed both classification rounds. Intrarater agreement was 'substantial' to 'almost perfect' for all joints ($0.64<\kappa<0.91$). Interrater agreement reached similar results ($0.63<\kappa<0.86$), except for the knee patterns during stance ($\kappa=0.49$, 'moderate agreement'). Experienced raters performed significantly better on patterns of the knee during stance and of the ankle during swing.

Conclusion

Apart from some knee patterns during stance, the results showed evidence that clinicians were able to use predefined joint patterns in CP with good confidence, even in case of limited experience with 3DGA.

Introduction

This study investigates the level of agreement with which clinicians can recognize specific joint patterns during gait in children with cerebral palsy (CP). Recently, a three-round Delphi consensus study defined a comprehensive set of joint patterns for ambulatory children with the spastic type of CP[1]. Supported by previous literature and the clinical expertise of the panel, three to seven patterns were defined for each of the different lower limb joints based on kinematic and kinetic data from three-dimensional gait analysis (3DGA). Increasing evidence is showing the importance of gait analysis for clinical decision making and suggests that a patient is more likely to have a better outcome post treatment if treatment is planned in accordance with recommendations from 3DGA, rather than if it is based solely on other clinical examinations[2–4]. Consequently, gait data has often been used in the past to classify the numerous gait abnormalities that characterize CP[5–8]. Yet, for any classification to be useful in clinical practice, it needs to be shown that clinicians can consistently assign the gait patterns in the CP population. Dobson et al. reported that reliability was evaluated for only two out of eighteen identified gait classifications in CP[5]. Rodda et al.[9] examined the inter-rater agreement of five gait patterns twice and found that six clinicians, experienced in 3DGA, reached overall acceptable agreement in 40 children, albeit with wide confidence intervals. Even the most frequently used and cited classification of Winters et al.[10], developed in 1987, was only tested for its reliability in 2006. Results showed that sixteen very experienced clinicians reached overall acceptable agreement on the four gait patterns in 34 children with CP when using kinematic data in combination with video data. However, a detailed study of the results indicated that agreement in two out of four patterns was poor[11]. Regarding the patterns defined in the aforementioned Delphi study[1], inter- and intrarater clinician agreement remains to be assessed. The present study used kinematic and kinetic data from a large sample of children with CP to measure the level of agreement on all different joint patterns in a group of clinicians with different (para)medical specialties and various levels of experience with 3DGA. The research questions were:

-Is there a good level of inter- and intrarater agreement for all joint patterns?

-Is experience with 3DGA a prerequisite for using the specified joint patterns during gait?

Methodology

Patient group

After the project was approved by the Medical Ethical Committee of University Hospitals Leuven (ref. s56036), a retrospective sample of convenience was recruited from the motion analysis laboratory of the hospital. Eligible patients were between 4 and 18 years old, classified as GMFCS level I-III, had predominantly spastic CP, and had undergone 3DGA. Previous treatments such as single-event multilevel surgery or botulinum toxin injections were allowed. All gait analyses were performed by experienced clinicians, using a ten to fifteen camera Vicon system (Vicon Motion Systems ltd., UK) and two AMTI force plates (Advanced Mechanical Technology Inc., USA), following the Plug-In-Gait marker model. Children walked barefoot at self-selected speed on a 10m walkway. Two representative gait trials of good quality were classified per side for each patient, resulting in 258 gait trials belonging to 92 patients with CP. For patients with unilateral CP, only the affected body side was classified.

Raters

No fixed amount of raters was targeted. In 2015, all participants of the gait courses preceding the conferences of the Gait & Clinical Movement Analysis Society (GCMAS) and the European Society for Movement Analysis in Adults and Children (ESMAC) were invited by email to take part in the study. Forty-two clinicians expressed an interest to participate and were asked to complete a short online survey, in which demographic information and data concerning their level of experience with 3DGA and CP was collected. In November 2015, a learning phase was proposed, where all candidates were encouraged via screencast presentations to get familiar with the online software that would be used to perform and collect the classifications (http://cmal-tools-leuven.be). Further presentations provided them with definitions and illustrations of the different joint gait patterns as they were presented in the Delphi study[1]. Pilot data from ten patients with CP, independent from the recruited study sample, were also available to all candidates to allow them to get acquainted with the software and with assigning the joint patterns. After the learning phase, 32 candidates confirmed their willingness to participate in the remainder of the study. They were divided in an ‘experienced’ (n=16) and ‘inexperienced’ (n=16) group based on their experience with 3DGA. Experience was measured by the raters’ self-reported level of expertise, years of

experience, and frequency of collecting or interpreting data of 3DGA (Table 1). The answers were scored on a five-point ordinal scale and then summed. The sixteen highest scoring raters formed the 'experienced' group. During the study, raters were unaware of the group they were allocated to.

The study comprised two classification rounds, each of them lasting five weeks with a minimal interval of one week in-between. During the second round, each rater received the same patients in a different order and with a different name. Raters were blinded to all patient information apart from the kinematic and kinetic reports which, in combination with a minimal one-week interval, maximized the independency between the ratings of the two rounds. Each rater received login details for www.cmal-tools-leuven.be, where they could find the reports and were able to make the classifications. Raters were also asked to indicate a trial as "unclassifiable" if they considered a patient not to fit any of the described patterns. Raters were free to classify their patients at times suitable for them, so it was possible to perform all classifications in one effort or to spread them over different times and days. To obtain a criterion classification, two expert raters who carried out the Delphi study classified all patient trials. In case of disagreement between these raters, consensus was sought with a member of the Delphi expert panel.

Statistical analysis

Power calculations were performed in R (v3.2.0; package 'kappaSize', MA Rotondi, https://cran.r-project.org/web/packages/kappaSize/index.html) and indicated that 82 patients should be classified by at least four raters to reach a kappa of 0.85 with confidence interval (CI) 0.75-0.95 (α=0.05; Table S1). Eighty-two patients were randomly selected from the 258 gait trials, but the presence of patterns with a prevalence of less than 10% was ensured in the selection. Because it was uncertain that all raters would finish two classification rounds, some margin was added in the distribution of patients among the raters. Each rater was therefore asked to classify 27 or 28 patients so that if no rater would drop out, each patient was to be classified at least five times in both the 'experienced' and 'inexperienced' group.

Inter- and intrarater agreement scores per joint were calculated for all raters and for the 'experienced' and 'inexperienced' group separately. Mean interrater agreement between the criterion classification and each rater was also calculated. For all interrater scores, four ratings were randomly selected per rater group if more were available. Trials that were reported as

'unclassifiable', were included in the analysis as the pattern of each joint that indicated 'No or minor gait deviations'. Fleiss' or Light's Kappa with 95% CIs and percentage of agreement (POA) were calculated in Matlab (Cardillo G. 2007, Cohen's kappa: http://www.mathworks.com/matlabcentral/fileexchange/15365)[12–14]. The strength of Kappa was interpreted as 'poor' ($\kappa<0$), 'slight' ($\kappa=0$ – 0.20), 'fair ' ($\kappa=0.21$-0.40), 'moderate' ($\kappa=0.41$-0.60), 'substantial' ($\kappa=0.61$-0.80), or 'almost perfect' agreement ($\kappa>0.80$)[15].

Results

The studied sample consisted of 82 children with spastic CP of which 47, 26, and 9 were classified as GMFCS I, II, and III, respectively. Out of the 82 children, 57 were male and 55 were bilaterally involved. Mean age was 9 years, 5 months (SD 3 years, 11 months), mean weight was 31.5 kg (SD 15.8kg) and mean height was 130.9cm (SD 24.0cm).

Table 1 describes the characteristics of the participating raters. Twenty-eight raters from 24 different clinical or research institutes across eight countries completed both classification rounds. Three raters dropped out (two inexperienced raters, one experienced rater), and one inexperienced rater only completed the first round. With this low drop-out number, at least four ratings were available for all 82 patients in each rater group (as required by power analysis, cfr. Statistical analysis). Only three out of 8701 ratings were missing in the first round and seven out of 8404 ratings in the second round. In the first round 4.7% of all ratings were deemed to be 'unclassifiable', slightly increasing up to 5.5% in the second round. Two raters accounted for one third of these unclassifiable ratings and the ratings were primarily assigned to the patterns of the hip in the sagittal (9.7%), coronal (13.0%), and transverse plane (8.5%). Table S2 presents an overview of the percentage of 'unclassifiable' ratings per rater and per joint. Most raters indicated that they chose the option 'unclassifiable' because they felt a pattern was missing, e.g. 'hyperextension during stance' for the hip in the sagittal plane.

Intrarater agreement

Table 2 provides intrarater agreement scores. POA for the entire rater group ranged from 82% to 94%, except for the knee during stance (POA=70%). Taking into account the possibility of agreement by chance, the raters reached 'substantial' to 'almost perfect' agreement for all joints, with the lowest scores for the knee patterns during stance ($\kappa=0.64$, CI=0.60-0.68) and the highest for the foot progression angle (FPA) ($\kappa=0.91$, CI=0.88-0.93).

Rater groups only differed significantly (no overlap in CI) for the knee patterns during stance (experienced raters, κ=0.70 vs. inexperienced raters, κ=0.57). Individual intrarater scores were characterized by much variability (Table S3-4); the mean kappa over all joints per rater was 0.76, with the experienced raters varying from 0.73 to 0.90, and the inexperienced raters from 0.59 to 0.89.

Table 1. Characteristics of rater groups.

	Experienced raters (n=15) (N)		**Inexperienced raters (n=14) (N)**	
Male / female	6 / 9		2 / 12	
Profession				
Pediatric orthopedic surgeon	6		1	
Physical therapist	6		8	
Kinesiologist	2		1	
(Pediatric) rehabilitation or neurology physician	1		3	
Clinical scientist in rehabilitation engineering	0		1	
Age				
≤ 30 years	2		3	
31-40 years	7		6	
41-50 years	6		3	
> 50 years	0		2	
	3DGA	**CP**	**3DGA**	**CP**
Self-reported expertise				
Fundamental awareness - basic knowledge	0	0	4	1
Novice - limited experience	0	0	7	2
Intermediate - practical application	9	7	3	5
Advanced - applied theory	6	6	0	5
Expert - recognized authority	0	2	0	1
Years of built-up experience				
< 2 years	4	0	12	2
2-5 years	7	6	2	8
6-10 years	3	4	0	1
11-15 years	0	2	0	0
> 15 years	1	3	0	3
Frequency of performing/interpreting 3DGA, Frequency of treating patients with CP				
Never	0	0	3	2
A few times per year	0	0	7	0
1-3 days per month	1	2	3	1
1-3 days per week	9	5	1	6
> 3 days per week	5	8	0	5

3DGA=three-dimensional gait analysis; CP=cerebral palsy.

Table 2. Intrarater agreement scores for different rater groups.

	All raters (n=28)			Experienced raters (n=15)			Inexperienced raters (n=13)		
	Kappa	**CI**	**POA (%)**	**Kappa**	**CI**	**POA (%)**	**Kappa**	**CI**	**POA (%)**
Sagittal plane									
Pelvis	0.76	0.73-0.80	83	0.75	0.70-0.81	82	0.78	0.72-0.83	84
Hip	0.71	0.67-0.75	82	0.74	0.69-0.80	84	0.67	0.60-0.74	79
Knee during stance	0.64	0.60-0.68	70	0.70	0.65-0.75	75	0.57	0.51-0.63	64
Knee during swing	0.80	0.76-0.83	83	0.79	0.74-0.83	83	0.80	0.76-0.85	84
Ankle during stance	0.77	0.74-0.81	83	0.76	0.71-0.81	82	0.79	0.74-0.84	84
Ankle during swing	0.74	0.71-0.78	82	0.75	0.70-0.80	82	0.74	0.68-0.79	81
Coronal plane									
Pelvis	0.81	0.77-0.84	86	0.82	0.77-0.86	87	0.79	0.74-0.85	85
Hip	0.71	0.67-0.75	82	0.74	0.68-0.80	84	0.68	0.61-0.74	79
Transverse plane									
Pelvis	0.77	0.74-0.81	84	0.78	0.73-0.83	84	0.77	0.72-0.82	83
Hip	0.87	0.83-0.90	92	0.90	0.86-0.94	94	0.83	0.78-0.88	90
Foot progression angle	0.91	0.88-0.93	94	0.93	0.89-0.96	95	0.89	0.85-0.93	93

CI=confidence interval; POA=average of each rater's percentage of agreement; Shaded cells indicate almost perfect agreement (κ>0.80). 'moderate agreement' (κ=0.41-0.60), 'substantial agreement' (κ=0.61-0.80), 'almost perfect agreement' (κ>0.80).

Interrater agreement

Table 3 presents all interrater agreement scores. POA varied from 74% to 91% for the rater group as a whole, except for the knee during stance (POA=58%). Taking into account agreement by chance, 'substantial' agreement was found for all joints but the knee during stance ('moderate' agreement, κ=0.49, CI=0.47-0.51) and the FPA ('almost perfect' agreement, κ=0.85, CI=0.83-0.89). The experienced rater group achieved significantly higher agreement scores (no overlap in CIs) for patterns of the knee during stance (κ=0.57 vs. κ=0.41), the ankle during swing (κ=0.76 vs. κ=0.51), and borderline higher agreement for the pelvis in the transverse plane (κ=0.78 vs. κ=0.66). The expert raters who classified all 82 patients reached 'almost perfect' agreement for all joints (0.80<κ<0.98), except for the knee during stance (κ=0.62, CI=0.50-0.74; Table S5).

Table 4 reports pattern-specific agreement for the experienced and inexperienced raters; Table S6 presents pattern-specific scores for all raters. Overall, six patterns reached notably lower levels of agreement compared to the other patterns of their respective joints: 'increased pelvic range of motion' (κ=0.47, CI=0.43-0.51), 'hip extension deficit' (κ=0.50, CI=0.45-0.54), 'delayed peak knee flexion in swing' (κ=0.57, CI=0.53-0.61), 'reversed second ankle rocker during stance' (κ=0.57, CI=0.53-0.61), 'insufficient ankle prepositioning in swing' (κ=0.49, CI=0.45-0.53), and 'excessive hip abduction in swing' (κ=0.55, CI=0.51-0.59). All but one of the knee patterns during stance reached low agreement results (all κ<0.60). Notably, the experienced raters agreed significantly more often than the inexperienced raters upon the patterns 'normal ankle during swing' (κ=0.75 vs. κ=0.50), 'insufficient prepositioning' (κ=0.69 vs. κ=0.34), and 'drop foot' (κ=0.78 vs. κ=0.53).

The mean interrater agreement between each rater and the criterion classification (Table 3) produced comparable results to the agreement levels of the rater group as a whole. The comparison with the criterion classifications scored significantly lower than the interrater agreement levels for the pelvis (κ=0.68 vs. κ=0.59) and knee patterns during swing (κ=0.72 vs. κ=0.66) in the sagittal plane, as well as the hip in the transverse plane (κ=0.78 vs. κ=0.69). Again, individual interrater comparisons with the expert classification varied highly between different raters (Table S7-8); mean kappa over all joints per rater was 0.66, ranging from 0.42 to 0.82.

Table 3. Interrater agreement scores between all raters (left side of table) and between raters and the criterion classification (right side of table).

	Interrater agreement scores									**Agreement between raters and criterion classification**								
	All raters (n=29)			**Experienced raters (n=15)**			**Inexperienced raters (n=14)**			**All raters (n=29)**			**Experienced raters (n=15)**			**Inexperienced raters (n=14)**		
	Kappa	**CI**	**POA (%)**	**Kappa**	**CI**	**POA (%)**	**Kappa**	**CI**	**POA (%)**	**Kappa**	**CI**	**POA (%)**	**Kappa**	**CI**	**POA (%)**	**Kappa**	**CI**	**POA (%)**
Sagittal plane																		
Pelvis	0.68	0.66-0.70	77	0.70	0.65-0.75	79	0.65	0.60-0.70	75	0.59	0.54-0.63	69	0.60	0.54-0.66	70	0.57	0.51-0.64	68
Hip	0.65	0.62-0.68	78	0.67	0.61-0.73	79	0.62	0.55-0.68	75	0.58	0.53-0.63	74	0.61	0.54-0.67	75	0.55	0.48-0.62	72
Knee ST	0.49	0.47-0.51	58	0.57	0.54-0.61	64	0.41	0.37-0.45	52	0.48	0.44-0.52	56	0.52	0.46-0.57	59	0.45	0.39-0.51	53
Knee SW	0.72	0.70-0.74	77	0.76	0.71-0.80	80	0.70	0.66-0.74	75	0.66	0.62-0.70	72	0.66	0.61-0.71	72	0.66	0.61-0.72	73
Ankle ST	0.68	0.66-0.70	76	0.70	0.65-0.75	77	0.68	0.63-0.72	75	0.69	0.65-0.73	77	0.68	0.63-0.74	77	0.70	0.65-0.76	78
Ankle SW	0.63	0.61-0.66	74	0.76	0.70-0.81	83	0.51	0.46-0.56	64	0.68	0.64-0.72	78	0.72	0.66-0.77	80	0.64	0.58-0.70	75
Coronal plane																		
Pelvis	0.71	0.69-0.74	79	0.73	0.67-0.78	81	0.69	0.64-0.75	78	0.70	0.66-0.74	79	0.70	0.64-0.75	79	0.70	0.64-0.76	79
Hip	0.66	0.64-0.69	78	0.67	0.62-0.73	79	0.65	0.60-0.70	77	0.62	0.57-0.67	77	0.64	0.57-0.70	79	0.60	0.53-0.67	75
Transverse plane																		
Pelvis	0.71	0.68-0.73	79	0.78	0.72-0.83	84	0.66	0.61-0.72	76	0.66	0.62-0.70	76	0.68	0.63-0.74	77	0.64	0.58-0.70	74
Hip	0.78	0.75-0.81	87	0.82	0.76-0.89	90	0.74	0.67-0.81	84	0.69	0.65-0.74	83	0.73	0.66-0.79	86	0.66	0.59-0.73	81
FPA	0.86	0.83-0.89	91	0.92	0.86-0.98	95	0.82	0.76-0.89	88	0.91	0.88-0.93	94	0.94	0.92-0.97	96	0.86	0.82-0.91	91

CI=confidence interval; POA=average percentage of agreement on each patient; ST=during stance; SW=during swing; FPA=foot progression angle; Shaded cells indicate almost perfect agreement (κ>0.80). 'moderate agreement' (κ=0.41-0.60), 'substantial agreement' (κ=0.61-0.80), 'almost perfect agreement' (κ>0.80).

Table 4. Comparison of pattern-specific agreement levels between experienced and inexperienced raters.

	Experienced raters (n=15)			Inexperienced raters		
	Kappa	**CI**	**POA(%)**	**Kappa**	**CI**	**POA(%)**
PELVIS in sagittal plane						
Normal pelvic motion/posture	0.83	0.74-0.92	86	0.75	0.66-0.84	80
Increased range of motion	0.51	0.42-0.60	55	0.39	0.31-0.48	44
Increased anterior tilt on average	0.61	0.52-0.70	69	0.60	0.52-0.69	71
Increased anterior tilt + increased range of motion	0.72	0.63-0.81	84	0.70	0.61-0.79	82
Decreased anterior tilt (posterior tilt)	0.89	0.80-0.97	89	0.74	0.66-0.83	75
Decreased anterior tilt (posterior tilt) + increased range of motion	0.74	0.66-0.83	75	0.66	0.57-0.75	67
HIP in sagittal plane						
Normal hip motion	0.71	0.62-0.80	85	0.69	0.60-0.78	83
Hip extension deficit	0.55	0.46-0.64	67	0.42	0.33-0.50	56
Continuous excessive hip flexion	0.74	0.65-0.83	80	0.70	0.61-0.79	79
KNEE during stance in sagittal plane						
Normal knee motion during stance	0.50	0.41-0.59	55	0.46	0.37-0.55	51
Increased knee flexion at initial contact	0.56	0.47-0.65	67	0.35	0.26-0.44	50
Increased knee flexion at initial contact + earlier knee extension movement	0.38	0.29-0.47	49	0.32	0.23-0.41	50
Knee hyperextension	0.69	0.61-0.78	73	0.72	0.63-0.81	75
Knee hyperextension + increased knee flexion at initial contact	0.68	0.59-0.77	70	0.45	0.36-0.54	48
Increased flexion in midstance + internal flexion moment present	0.60	0.51-0.69	67	0.34	0.25-0.43	44
Increased flexion in midstance + internal extension moment present	0.75	0.66-0.83	78	0.46	0.37-0.55	51
KNEE during swing in sagittal plane						
Normal knee motion during swing	0.80	0.72-0.89	86	0.72	0.64-0.81	80
Delayed peak knee flexion	0.58	0.49-0.67	65	0.59	0.51-0.68	67
Increased peak knee flexion	0.88	0.79-0.97	89	0.74	0.65-0.83	77
Increased + delayed peak knee flexion	0.80	0.71-0.88	82	0.77	0.68-0.86	80
Decreased peak knee flexion	0.78	0.69-0.86	80	0.72	0.63-0.81	75
Decreased + delayed peak knee flexion	0.71	0.62-0.80	76	0.67	0.58-0.76	73

Table 4. Continued.

	Experienced Raters (n=15)			Inexperienced raters (n=14)		
	Kappa	**CI**	**POA(%)**	**Kappa**	**CI**	**POA(%)**
ANKLE during stance in sagittal plane						
Normal ankle motion during stance	0.75	0.66-0.84	82	0.71	0.62-0.80	80
Horizontal second ankle rocker	0.63	0.54-0.72	77	0.68	0.59-0.77	79
Reversed second ankle rocker	0.60	0.51-0.69	65	0.56	0.47-0.65	62
Equinus gait	0.74	0.65-0.82	76	0.73	0.64-0.82	76
Calcaneus gait	0.81	0.72-0.90	84	0.67	0.59-0.76	72
ANKLE during swing in sagittal plane						
Normal ankle motion during swing	0.75	0.66-0.83	85	0.50	0.42-0.59	69
Insufficient prepositioning in terminal swing	0.69	0.60-0.78	73	0.34	0.25-0.42	44
Continuous plantarflexion during swing (drop foot)	0.78	0.70-0.87	85	0.53	0.45-0.62	65
Excessive dorsiflexion during swing	0.80	0.71-0.89	83	0.63	0.54-0.72	71
PELVIS in coronal plane						
Normal pelvic motion/posture	0.68	0.59-0.77	80	0.65	0.56-0.74	77
Increased pelvic range of motion	0.70	0.61-0.79	79	0.70	0.61-0.79	79
Continuous pelvic elevation	0.76	0.67-0.84	78	0.67	0.58-0.76	70
Continuous pelvic depression	0.81	0.72-0.90	86	0.75	0.66-0.84	81
HIP in coronal plane						
Normal hip motion	0.64	0.55-0.73	84	0.68	0.59-0.77	84
Excessive hip abduction in swing	0.56	0.47-0.65	63	0.56	0.48-0.65	64
Continuous excessive hip abduction	0.74	0.66-0.83	78	0.64	0.55-0.73	69
Continuous excessive hip adduction	0.76	0.67-0.85	79	0.70	0.61-0.78	74
PELVIS in transverse plane						
Normal pelvic motion/posture	0.77	0.69-0.86	85	0.59	0.50-0.67	71
Increased pelvic range of motion	0.69	0.60-0.78	78	0.59	0.50-0.68	71
Excessive pelvic external rotation during the gait cycle	0.81	0.73-0.90	87	0.72	0.64-0.81	81
Excessive pelvic internal rotation during the gait cycle	0.93	0.84-1.00	93	0.86	0.77-0.95	88

Table 4. Continued.

	Experienced Raters (n=15)			Inexperienced raters (n=14)		
	Kappa	**CI**	**POA(%)**	**Kappa**	**CI**	**POA(%)**
HIP in transverse plane						
Normal hip motion	0.81	0.72-0.90	92	0.71	0.62-0.80	86
Excessive hip external rotation during the gait cycle	0.84	0.75-0.93	86	0.77	0.68-0.86	80
Excessive hip internal rotation during the gait cycle	0.84	0.75-0.93	88	0.76	0.67-0.85	84
FOOT progression angle						
Normal foot progression angle	0.89	0.80-0.98	94	0.78	0.69-0.87	86
Outtoeing	0.98	0.90-1.00	99	0.85	0.76-0.94	90
Intoeing	0.89	0.80-0.98	92	0.84	0.75-0.93	89

CI=confidence interval; POA= average percentage of agreement on each patient; Shaded cells indicate fair or moderate agreement ($\kappa<0.60$). 'fair agreement' ($\kappa=0.21$-0.40) 'moderate agreement' ($\kappa=0.41$-0.60), 'substantial agreement' ($\kappa=0.61$-0.80), 'almost perfect agreement' ($\kappa>0.80$). A more detailed description of the patterns is available in Nieuwenhuys et al.[1].

Interrater agreement scores between the first and second classification round were almost equal; differences in kappa values between rounds were within 0.03, except for the agreement for the knee during stance and hip in the sagittal plane, with kappa-increases of 0.05.

Discussion

In general, this study found that the level of agreement with which clinicians can recognize recently specified joint patterns during gait in children with spastic CP was 'good' for the patterns of all joints across all planes, except for the knee during stance in the sagittal plane ('moderate agreement'). Significantly higher agreement levels in the experienced rater group for the classification of knee patterns during stance and ankle patterns during swing, suggest that experience with 3DGA might be advantageous in assigning the patterns reliably.

Intrarater agreement is identified as being 'almost perfect' for the patterns of the FPA and 'substantial' for the patterns of all other joints. Intrarater agreement of gait patterns in CP has been reported by Rodda et al.[9], who found agreement of six experienced raters for five sagittal gait patterns in children with diplegia to be varying between κ=0.66 to κ=1 based on sagittal plane kinematics and video recordings. For the same patterns, Stott et al.[16] found the intrarater agreement of five raters to range between κ=0.50 to κ=0.68 based on video recordings. In the present study individual intrarater agreement scores were similar (mean kappa over all joints within 0.59 to 0.90) and without marked differences between experienced and inexperienced raters. It appears that intrarater agreement on gait patterns in children with CP cannot easily be generalized to all clinicians but is rather rater-specific.

Interrater agreement levels and POA in both classification rounds indicated good agreement for most joints, apart from the FPA ('almost perfect' agreement) and the knee during stance ('moderate' agreement). Both expert raters also achieved notably lower agreement for the knee during stance. Given that the knee during stance is characterized by the highest number of patterns (n=7), it also has the highest probability of disagreement. Consequently, it is not surprising that lower agreement results were found. Nonetheless, with POA=58% and κ=0.49 for all raters, it is clear that these patterns cannot yet be used by clinicians with confidence. The same conclusion applies to the six other patterns that reached relatively low agreement scores (κ<0.60). It is not uncommon to find an overall 'good' (to 'very good') level of agreement for a classification of gait in CP and at the same time, discover less acceptable levels of agreement for a few specific patterns. The present study identified 12 out of 49

patterns (25%) with ‘fair’ or ‘moderate’ agreement. Rodda et al.[9] found similar results for the patterns ‘jump gait’ and ‘apparent equinus’ (two out of five patterns, 40%), and likewise, Dobson et al.[11] found this result for group II and III (two out of four patterns, 50%) of the Winters’ classification[10]. These results suggest that current definitions for these patterns might not be descriptive (or restrictive) enough and adaptations could be necessary.

The experienced rater group consistently reached slightly higher agreement levels (average kappa difference of 0.09 between both rater groups for interrater agreement). Given that the patterns evaluated in this study were defined by a very experienced panel, it might be hypothesized that the patterns therefore fit more naturally with raters who are already more familiar with 3DGA. This is contrary to the findings of Dobson et al. who found slightly higher agreement for physiotherapists with less experience in 3DGA, hypothesizing that clinicians with more experience in 3DGA might be more lenient to individual interpretations of the patterns[11]. It should be noted that the rater group in that study had much more experience with 3DGA than both rater groups in the present study. Even with slightly lower agreement results, the inexperienced rater group still reached substantial to almost perfect agreement results for all joints, except for the knee during stance and ankle during swing. Perhaps inexperienced raters might benefit from a more controlled learning phase.

A few limitations should be discussed. Firstly, 4.7% percent of all ratings for all patterns were deemed to be unclassifiable, and the majority of these ratings were made for patterns of the hip across the three anatomical planes. Compared to research on the patterns by Winters et al.[10], this is a relatively low number, which could be interpreted as an indirect confirmation of the content validity of the patterns[17,18]. Future research should therefore further examine the causes and characteristics of the unclassifiable hip trials. By conservatively considering unclassifiable trials as ‘normal or minor gait deviations’ in the analysis, somewhat different agreement results might have been found if these trials were considered as a separate group. Secondly, the effect of prevalence or bias on kappa was not examined. The penalizing effect of asymmetrical prevalence of joint patterns on kappa did however not appear to be a problem when comparing kappa to POA. Unfortunately, statistical procedures that adjust kappa to minimize these effects such as prevalence-adjusted bias-adjusted kappa cannot be used in the case of more than two classes or two raters[19,20]. The learning phase of the present study was intentionally limited and uncontrolled, and there was no assessment whether all raters consulted the screencast presentations or whether they classified the pilot data before the first

classification round. This could explain why slightly lower agreement results were found for the interrater agreement between the criterion classification and the rater groups compared to the overall interrater agreement results.

In conclusion, the present study shows promise that even with a limited learning phase or limited experience with 3DGA, clinicians will be able to assign most joint patterns during gait confidently. However, future research should examine characteristics of unclassifiable patients and re-examine the definitions of specific patterns with low agreement. It is not unreasonable to assume that agreement might significantly improve given a stricter learning phase, where for instance raters are instructed to classify pilot patients with CP after which feedback on the results is provided. This is the first study using 3DGA data to evaluate the level of clinician agreement on gait patterns in CP, in a large international rater group from various clinical and research centers, with a range of clinical professions, and with different levels of experience with 3DGA and CP. By keeping the learning phase intentionally brief and uncontrolled, the generalizability of the presented results is maximized.

Acknowledgements

The authors would like to express their gratitude to all participants who devoted valuable time and effort in learning and applying the classification of joint patterns during gait in children with cerebral palsy. More specifically, we thank Ms. S Aydýl, Ms. E Loof, Mr. MV Belthur, Ms. P Willmarth, Mr. F Krüger, Ms. S Mader, Mr. L Becerra, Ms. K Kaiser, Mr. GH Ortiz, Ms. C Weigel, Mr. L de Oliveira, Ms. A Perry, Mr. K Beng, Ms. R Vance, Mr. LN Veilleux, Ms. A Hilderley (Holland Bloorview Kids Rehabilitation Hospital), and Ms. K Dieruf and Ms. M Barkocy (University of New Mexico, division of physical therapy).

A. Nieuwenhuys is supported by an OT project of KU Leuven University (OT/12/100). E. Papageorgiou is supported by the MD Paedigree project, a Model-Driven Paediatric European Digital Repository, partially funded by the European Commission under FP7 – ICT Programme (grant agreement no: 600932, http://www.md-paedigree.eu). Funders had no involvement in study design, data collection, data analysis, manuscript preparation and / or decision to publish.

References

1. Nieuwenhuys A, Õunpuu S, Van Campenhout A, et al. Identification of joint patterns during gait in children with cerebral palsy: A Delphi consensus study. *Dev Med Child Neurol*. 2016;58(3):306-313.
2. Wren TAL, Otsuka NY, Bowen RE, et al. Outcomes of lower extremity orthopedic surgery in ambulatory children with cerebral palsy with and without gait analysis: Results of a randomized controlled trial. *Gait Posture*. 2013;38(2):236-241.
3. Lofterød B, Terjesen T. Results of treatment when orthopaedic surgeons follow gait-analysis recommendations in children with CP. *Dev Med Child Neurol*. 2008;50(7):503-509.
4. Cook RE, Schneider I, Hazlewood ME, Hillman SJ, Robb JE. Gait Analysis Alters Decision-Making in Cerebral Palsy. *J Pediatr Orthop*. 2003:292-295.
5. Dobson F, Morris ME, Baker R, Graham HK. Gait classification in children with cerebral palsy: A systematic review. *Gait Posture*. 2007;25(1):140-152.
6. Carriero A, Zavatsky A, Stebbins J, Theologis T, Shefelbine SJ. Determination of gait patterns in children with spastic diplegic cerebral palsy using principal components. *Gait Posture*. 2009;29(1):71-75.
7. Ferrari A, Alboresi S, Muzzini S, Pascale R, Perazza S, Cioni G. The term diplegia should be enhanced. Part I: a new rehabilitation oriented classification of cerebral palsy. *Eur J Phys Rehabil Med*. 2008;44(2):195-201.
8. Toro B, Nester CJ, Farren PC. Cluster analysis for the extraction of sagittal gait patterns in children with cerebral palsy. *Gait Posture*. 2007;25(2):157-165.
9. Rodda JM, Graham HK, Carson L, Galea MP, Wolfe R. Sagittal gait patterns in spastic diplegia. *J Bone Joint Surg Br*. 2004;86(2):251-258.
10. Winters TF, Gage JR, Hicks R. Gait patterns in spastic hemiplegia in children and young adults. *J Bone Joint Surg Am*. 1987;69:437-441. doi:10.1302/0301-620X.86B2.13878.
11. Dobson F, Morris M, Baker R, Wolfe R, Graham H. Clinician agreement on gait pattern ratings in children with spastic hemiplegia. *Dev Med Child Neurol*. 2006;48(6):429-435.
12. Fleiss JL. Measuring nominal scale agreement among many raters. *Psychol Bull*. 1971;76(5):378-382.
13. Light RJ. Measures of response agreement for qualitative data: Some generalizations and alternatives. *Psychol Bull*. 1971;76(5):365-377.
14. Cohen J. A coefficient of agreement for nominal scales. *Educ Psychol Meas*. 1960;20(1):37-46.
15. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics*. 1977;33(1):159-174.
16. Stott NS, Atherton WG, Mackey AH, Galley IJ, Nicol RO, Walsh SJ. The reliability and validity of assessment of sagittal plane deviations in children who have spastic diplegia. *Arch Phys Med Rehabil*. 2005;86(12):2337-2341.
17. Riad J, Haglund-Akerlind Y, Miller F. Classification of Spastic Hemiplegic Cerebral Palsy in Children. *J Pediatr Orthop*. 2007;27(7):758-764.
18. McDowell BC, Kerr C, Kelly C, Salazar J, Cosgrove, A. The validity of an existing gait classification system when applied to a representative population of children with hemiplegia. *Gait Posture*. 2008;28(3):442-447.
19. Sim J, Sim J, Wright CC, Wright CC. The kappa statistic in reliability studies: use,

interpretation, and sample size requirements. *Phys Ther*. 2005;85(3):257-268.

20. Byrt T, Bishop J, Carlin JB. Bias, prevalence and kappa. *J Clin Epidemiol*. 1993;46(5):423-429.

Supporting Information

Table S1. Sample size estimations indicated that a sample size of 82 patients needed to be classified by at least four raters to ensure sufficient power of the study.

	Number of raters			
	n=3	**n=4**	**n=5**	**n=6**
Sagittal plane				
- pelvis (6 patterns)	---	---	---	---
- hip (3 patterns)	82	65	56	50
- knee during stance (7 patterns)	---	---	---	---
- knee during swing (6 patterns)	---	---	---	---
- ankle during stance (5 patterns)	---	---	---	---
- ankle during swing (4 patterns)	77	61	53	48
Coronal plane				
- pelvis (4 patterns)	73	58	50	45
- hip (4 patterns)	104	82	71	64
Transverse plane				
- pelvis (4 patterns)	72	57	50	45
- hip (3 patterns)	104	82	71	63
- foot progression angle (3 patterns)	86	68	59	53

Numbers in the table indicate the required number of patients to be classified by a predefined number of raters to reach a kappa of 0.85 with confidence interval 0.75-0.95 at an α-level of 0.05. The appraised prevalence of each of the patterns was also taken into account. An estimate of kappa = 0.85 was based on a pilot study (Nieuwenhuys et al., doi:10.1016/j.gaitpost.2015.06.088). Because required sample sizes are highest with a lower number of patterns, power calculations were not performed for joints with more than 4 patterns. The largest sample size estimations were found for the hip in the coronal and transverse plane because the prevalence of the patterns in these joints was very asymmetrically distributed. After the final number of participating raters was clear and an estimate could be made on the expected time investment per rater, it was decided to choose a sample size of 82 patients which needed to be classified by 4 raters.

Table S2. Overview of unclassifiable ratings (%) in the first round per rater and per joint. In total, 4.7 % of all ratings were rated unclassifiable.

	Sagittal plane						Coronal plane		Transverse plane			Total per	Ratings per
	Pelvis	Hip	Knee	Knee SW	Ankle	Ankle	Pelvis	Hip	Pelvis	Hip	FPA	rater (%)	rater (N)
Rater 1	-	0.89	-	-	0.13	-	0.13	0.38	0.13	0.25	0.13	**5.4**	297
Rater 2	-	0.63	-	0.13	-	-	-	0.76	0.38	0.25	-	**5.5**	308
Rater 3	-	0.63	-	0.38	0.25	0.13	0.13	0.38	0.13	0.13	-	**5.7**	297
Rater 5	-	0.13	-	-	0.38	0.25	0.25	0.51	-	-	-	**4.0**	297
Rater 6	-	0.13	-	0.38	0.25	0.13	0.13	0.13	-	0.13	-	**3.4**	297
Rater 8	-	0.13	-	-	0.13	-	0.25	-	-	0.25	-	**2.0**	296
Rater 10	-	0.13	-	0.25	-	-	-	-	0.13	0.13	-	**1.7**	297
Rater 16	-	1.27	0.38	0.51	0.13	1.26	-	2.15	-	1.01	0.76	**19.1**	308
Rater 17	-	0.38	-	-	-	-	0.25	1.14	-	0.13	-	**5.1**	297
Rater 18	-	0.13	-	-	-	-	-	-	-	-	-	**0**	296
Rater 19	-	0.51	0.51	0.25	0.51	0.51	0.51	0.76	-	0.38	0.13	**10.8**	297
Rater 22	-	0.38	0.13	0.13	0.25	-	-	0.25	-	0.13	-	**3.4**	297
Rater 25	-	0.13	-	-	0.13	-	-	0.63	-	-	-	**2.3**	297
Rater 27	-	0.38	-	-	-	-	0.13	0.13	-	-	-	**1.6**	308
Rater 29	-	-	-	-	-	-	-	-	-	-	-	**0**	308
Rater 4	-	0.25	-	-	-	-	-	0.13	-	0.13	-	**1.4**	297
Rater 7	-	-	-	-	-	-	-	-	-	-	-	**0**	297
Rater 9	-	0.51	-	0.13	-	-	-	0.63	-	0.38	0.13	**4.5**	308
Rater 11	-	0.13	-	-	-	0.25	0.13	0.13	-	-	-	**1.7**	297
Rater 12	-	-	-	-	-	-	-	-	-	-	-	**0**	297
Rater 13	-	0.38	-	-	-	-	0.13	0.76	0.13	1.39	-	**7.4**	297
Rater 14	-	-	-	-	-	-	-	-	-	-	-	**0**	297
Rater 15	-	-	-	-	-	-	-	-	-	-	-	**0**	307
Rater 20	-	0.63	-	0.13	0.13	0.38	0.25	0.63	0.25	0.25	0.25	**7.8**	297
Rater 21	-	-	-	-	-	-	-	-	-	-	-	**0**	308
Rater 23	-	1.27	0.51	0.38	0.63	1.52	0.38	1.64	1.01	0.88	0.25	**22.6**	297
Rater 24	0.25	0.38	0.25	-	-	-	0.76	1.26	0.88	1.64	-	**14.5**	297
Rater 26	-	-	-	-	-	-	0.13	-	-	0.13	-	**0.7**	308
Rater 28	-	0.38	-	0.13	0.13	0.25	0.13	0.63	-	0.88	0.13	**7.1**	297
Total per joint (%)	**0.3**	**9.7**	**1.8**	**2.8**	**3.0**	**4.7**	**3.7**	**13.0**	**3.0**	**8.5**	**1.8**		
Ratings per joint (N)	791	790	790	791	791	791	790	791	791	791	791		

FPA=foot progression angle; ST=during stance; SW=during swing; - = no trials indicated as unclassifiable; Experienced rater group is shaded.

Table S3. Overview of intrarater agreement kappa and POA for each rater in the inexperienced rater group (N=13).

		Sagittal plane						Coronal plane		Transverse plane			
		Pelvis	**Hip**	**Knee ST**	**Knee SW**	**Ankle ST**	**Ankle SW**	**Pelvis**	**Hip**	**Pelvis**	**Hip**	**FPA**	**Mean per rater**
Rater 4	**Kappa**	0.88	0.87	0.73	0.84	0.95	0.84	0.95	0.84	0.88	0.94	0.73	0.86
	POA (%)	93	93	78	89	96	89	96	89	93	96	81	90
Rater 7	**Kappa**	0.69	0.76	0.78	0.95	0.74	0.73	0.65	0.66	0.84	1.00	1.00	0.80
	POA (%)	78	85	81	96	81	81	78	78	89	100	100	86
Rater 9	**Kappa**	0.59	0.65	0.54	0.66	0.52	0.51	0.90	0.42	0.75	0.75	0.72	0.64
	POA (%)	82	79	64	71	64	68	93	64	82	86	82	76
Rater 11	**Kappa**	0.79	0.50	0.43	0.85	0.84	0.83	0.82	0.48	0.73	0.86	0.94	0.73
	POA (%)	85	67	59	89	89	89	89	70	81	93	96	82
Rater 12	**Kappa**	0.89	1.00	0.68	0.86	0.85	0.78	0.89	0.72	0.79	0.81	1.00	0.84
	POA (%)	93	100	74	89	89	85	93	85	85	89	100	89
Rater 13	**Kappa**	0.94	0.78	0.87	0.86	0.86	0.95	0.95	0.85	0.84	1.00	0.94	0.89
	POA (%)	96	85	89	89	89	96	96	89	89	100	96	92
Rater 15	**Kappa**	0.90	0.82	0.33	0.96	0.85	0.84	0.73	0.88	0.79	0.81	1.00	0.81
	POA (%)	93	89	48	96	89	89	82	93	86	89	100	87
Rater 20	**Kappa**	0.85	0.46	0.64	0.78	0.89	0.75	0.89	0.76	0.89	0.94	0.89	0.79
	POA (%)	89	63	70	81	93	81	93	85	93	96	93	85
Rater 21	**Kappa**	0.64	0.50	0.38	0.61	0.56	0.48	0.54	0.71	0.66	0.58	0.78	0.59
	POA (%)	75	68	57	68	68	61	68	82	75	71	85	71
Rater 23	**Kappa**	0.57	0.74	0.49	0.81	0.90	0.45	0.60	0.28	0.52	0.73	0.78	0.62
	POA (%)	70	85	63	85	93	63	70	59	63	81	85	74
Rater 24	**Kappa**	0.70	0.47	0.51	0.71	0.90	0.84	0.83	0.79	0.76	0.91	0.88	0.75
	POA (%)	78	65	62	78	93	89	88	85	85	96	93	83
Rater 26	**Kappa**	0.73	0.65	0.42	0.73	0.67	0.73	0.64	0.73	0.68	0.57	0.89	0.68
	POA (%)	82	82	54	79	75	82	75	82	79	79	93	78
Rater 28	**Kappa**	0.71	0.51	0.24	0.74	0.62	0.72	0.84	0.50	0.78	0.88	1.00	0.69
	POA (%)	78	67	37	81	74	81	89	67	85	93	100	77
Mean per joint	**Kappa**	0.76	0.67	0.54	0.80	0.78	0.73	0.79	0.66	0.76	0.83	0.89	
	POA (%)	84	79	64	84	84	81	85	79	83	90	93	

FPA=foot progression angle; ST=during stance; SW=during swing; POA=percentage of agreement.

Table S4. Overview of intrarater agreement kappa and POA for each rater in the experienced rater group (N=15).

		Sagittal plane						Coronal plane		Transverse plane			
		Pelvis	**Hip**	**Knee ST**	**Knee SW**	**Ankle ST**	**Ankle SW**	**Pelvis**	**Hip**	**Pelvis**	**Hip**	**FPA**	**Mean per rater**
Rater 1	**Kappa**	0.72	0.64	0.74	0.69	0.66	0.57	0.84	0.77	0.85	0.94	0.60	0.73
	POA (%)	81	81	81	74	74	70	89	89	89	96	74	82
Rater 2	**Kappa**	0.65	0.64	0.56	0.82	0.67	0.57	0.85	0.84	0.63	0.83	1.00	0.73
	POA (%)	75	79	64	86	75	71	89	89	75	89	100	81
Rater 3	**Kappa**	0.79	0.77	0.69	0.73	0.56	0.57	0.83	0.41	0.71	0.93	1.00	0.73
	POA (%)	85	85	74	78	67	74	89	63	81	96	100	81
Rater 5	**Kappa**	0.57	0.72	0.69	0.85	0.78	0.90	0.80	0.68	0.64	0.61	0.94	0.74
	POA (%)	67	85	74	89	85	93	85	78	74	78	96	82
Rater 6	**Kappa**	1.00	0.83	0.76	0.85	0.90	0.79	0.95	0.95	0.89	1.00	0.94	0.90
	POA (%)	100	89	81	89	93	85	96	96	93	100	96	93
Rater 8	**Kappa**	0.83	0.67	0.77	0.68	0.76	0.49	0.63	0.71	0.74	0.92	0.94	0.74
	POA (%)	89	81	81	74	81	70	78	81	81	96	96	83
Rater 10	**Kappa**	0.80	0.61	0.61	0.85	0.82	0.70	0.89	0.62	0.74	0.81	0.89	0.76
	POA (%)	85	74	70	89	89	78	92	74	81	89	93	83
Rater 16	**Kappa**	0.69	0.81	0.81	0.72	0.60	0.89	0.72	0.57	0.67	0.89	0.89	0.75
	POA (%)	79	89	86	79	75	93	86	82	79	93	93	85
Rater 17	**Kappa**	0.84	0.72	0.68	0.76	0.74	0.70	0.57	0.88	0.78	1.00	0.94	0.78
	POA (%)	89	81	74	81	81	78	70	93	85	100	96	85
Rater 18	**Kappa**	0.65	0.77	0.82	0.73	0.80	0.90	0.95	0.84	0.79	0.87	0.89	0.82
	POA (%)	78	85	85	78	85	93	96	89	85	93	93	87
Rater 19	**Kappa**	0.74	0.76	0.42	0.81	0.84	0.79	0.63	0.61	0.79	0.87	0.88	0.74
	POA (%)	81	85	52	85	89	85	74	74	85	93	93	81
Rater 22	**Kappa**	0.83	0.82	0.47	0.72	0.67	0.61	0.69	0.77	0.80	0.88	1.00	0.75
	POA (%)	89	89	59	78	74	74	78	89	85	93	100	82
Rater 25	**Kappa**	0.42	0.68	0.56	0.82	0.71	0.79	0.84	0.58	0.72	0.93	1.00	0.73
	POA (%)	59	81	63	85	78	85	89	74	81	96	100	81
Rater 27	**Kappa**	0.95	0.71	0.87	0.87	0.84	0.89	1.00	0.78	0.89	1.00	1.00	0.89
	POA (%)	96	86	89	89	89	93	100	86	93	100	100	93
Rater 29	**Kappa**	0.69	0.76	0.87	0.87	0.95	0.87	0.89	0.94	0.90	1.00	0.94	0.88
	POA (%)	82	86	89	89	96	93	93	96	93	100	96	92
Mean per joint	**Kappa**	0.74	0.73	0.69	0.79	0.75	0.74	0.81	0.73	0.77	0.90	0.92	
	POA (%)	82	84	75	83	82	82	87	84	84	94	95	

FPA=foot progression angle; ST=during stance; SW=during swing; POA=percentage of agreement.

Table S5. Agreement scores for two expert raters on 82 patients with cerebral palsy.

	Expert raters (n=2)		
	Kappa	**CI**	**POA (%)**
Sagittal plane patterns			
Pelvis (n=6)	0.80	0.70-0.91	85
Hip (n=3)	0.90	0.81-0.98	94
Knee during stance (n=7)	0.62	0.50-0.74	68
Knee during swing (n=6)	0.88	0.80-0.96	90
Ankle during stance (n=5)	0.83	0.74-0.93	88
Ankle during swing (n=4)	0.81	0.70-0.91	87
Coronal plane patterns			
Pelvis (n=4)	0.96	0.91-1.00	98
Hip (n=4)	0.85	0.74-0.96	91
Transverse plane patterns			
Pelvis (n=4)	0.98	0.95-1.00	99
Hip (n=3)	0.92	0.84-1.00	96
Foot progression angle (n=3)	0.92	0.85-0.99	95

CI=confidence interval; POA= average percentage of agreement on each patient. Shaded cells indicate almost perfect agreement (κ>0.80).

Table S6. Pattern-specific POA and kappa for the entire rater group.

	All raters (n=29)		
	Kappa	**CI**	**POA (%)**
Pelvis			
Normal pelvic motion/posture	0.80	0.76-0.84	84
Increased range of motion	0.47	0.43-0.51	52
Increased anterior tilt on average	0.61	0.57-0.65	70
Increased anterior tilt + increased range of motion	0.70	0.66-0.74	83
Decreased anterior tilt (posterior tilt)	0.84	0.80-0.88	84
Decreased anterior tilt (posterior tilt) + increased range of motion	0.72	0.68-0.76	72
Hip			
Normal hip motion	0.72	0.68-0.76	85
Hip extension deficit	0.50	0.45-0.54	63
Continuous excessive hip flexion	0.71	0.66-0.75	78
Knee during stance phase			
Normal knee motion during stance	0.50	0.46-0.54	55
Increased knee flexion at initial contact	0.47	0.43-0.51	60
Increased knee flexion at initial contact + earlier knee extension movement	0.33	0.29-0.37	49
Knee hyperextension	0.71	0.67-0.75	74
Knee hyperextension + increased knee flexion at initial contact	0.54	0.50-0.58	57
Increased flexion in midstance + internal flexion moment present	0.47	0.43-0.51	56
Increased flexion in midstance + internal extension moment present	0.59	0.54-0.63	63
Knee during swing phase			
Normal knee motion during swing	0.76	0.72-0.80	83
Delayed peak knee flexion	0.57	0.53-0.61	65
Increased peak knee flexion	0.77	0.73-0.81	80
Increased + delayed peak knee flexion	0.77	0.72-0.81	80
Decreased peak knee flexion	0.75	0.71-0.79	78
Decreased + delayed peak knee flexion	0.70	0.66-0.74	75
Ankle during stance phase			
Normal ankle motion during stance	0.73	0.69-0.77	81
Horizontal second ankle rocker	0.65	0.61-0.69	77
Reversed second ankle rocker	0.57	0.53-0.61	63
Equinus gait	0.73	0.69-0.77	76
Calcaneus gait	0.72	0.68-0.76	76
Ankle during swing phase			
Normal ankle motion during swing	0.63	0.59-0.68	78
Insufficient prepositioning in terminal swing	0.49	0.45-0.53	57
Continuous plantarflexion during swing (drop foot)	0.65	0.61-0.69	75
Excessive dorsiflexion during swing	0.73	0.69-0.77	78

Table S6. Continued.

	All raters (n=29)		
	Kappa	**CI**	**POA (%)**
Pelvis in coronal plane			
Normal pelvic motion/posture	0.67	0.63-0.71	78
Increased pelvic range of motion	0.71	0.66-0.75	79
Continuous pelvic elevation	0.69	0.65-0.73	72
Continuous pelvic depression	0.79	0.75-0.83	84
Hip in coronal plane			
Normal hip motion	0.67	0.62-0.71	84
Excessive hip abduction in swing	0.55	0.51-0.59	63
Continuous excessive hip abduction	0.71	0.67-0.75	76
Continuous excessive hip adduction	0.74	0.70-0.78	78
Pelvis in transverse plane			
Normal pelvic motion/posture	0.67	0.63-0.71	77
Increased pelvic range of motion	0.61	0.57-0.65	72
Excessive pelvic external rotation during the gait cycle	0.77	0.73-0.81	84
Excessive pelvic internal rotation during the gait cycle	0.89	0.85-0.93	90
Hip in transverse plane			
Normal hip motion	0.76	0.71-0.80	89
Excessive hip external rotation during the gait cycle	0.80	0.76-0.84	83
Excessive hip internal rotation during the gait cycle	0.80	0.76-0.84	86
Foot progression angle			
Normal foot progression angle	0.82	0.78-0.86	89
Outtoeing	0.89	0.85-0.94	92
Intoeing	0.86	0.82-0.90	90

CI=confidence interval; POA= average percentage of agreement on each patient; Shaded cells indicate almost perfect agreement (κ>0.80). 'moderate agreement' (κ=0.41-0.6), 'substantial agreement' (κ=0.61-0.8), or 'almost perfect agreement' (κ=0.81-100).

Table S7. Interrater agreement kappa and POA for each inexperienced rater vs. criterion classification (N=14).

		Sagittal plane						Coronal plane		Transverse plane			
		Pelvis	**Hip**	**Knee ST**	**Knee SW**	**Ankle ST**	**Ankle SW**	**Pelvis**	**Hip**	**Pelvis**	**Hip**	**FPA**	**Mean per rater**
Rater 4	**Kappa**	0.77	0.59	0.74	0.70	0.85	0.84	0.53	0.89	0.42	0.54	0.73	0.69
	POA (%)	85	74	78	78	89	89	67	93	59	74	81	79
Rater 7	**Kappa**	0.56	0.59	0.70	0.81	0.64	0.66	0.71	0.65	0.84	0.85	1.00	0.73
	POA (%)	67	74	74	85	74	78	81	78	89	93	100	81
Rater 9	**Kappa**	0.46	0.50	0.30	0.57	0.52	0.68	0.70	0.48	0.52	0.32	0.79	0.53
	POA (%)	68	68	39	64	64	79	79	68	64	57	86	67
Rater 11	**Kappa**	0.41	0.47	0.30	0.75	0.79	0.77	1.00	0.60	0.74	0.85	0.94	0.69
	POA (%)	52	67	41	81	85	85	100	78	81	93	96	78
Rater 12	**Kappa**	0.63	0.64	0.52	0.64	0.75	0.50	0.89	0.59	0.73	0.93	0.94	0.71
	POA (%)	74	81	59	70	81	67	93	81	81	96	96	80
Rater 13	**Kappa**	0.83	0.83	0.73	0.77	0.95	0.80	0.89	0.78	0.74	0.79	0.94	0.82
	POA (%)	89	89	78	81	96	85	93	85	81	89	96	88
Rater 14	**Kappa**	0.26	0.32	0.30	0.56	0.48	0.46	0.49	0.30	0.43	0.75	0.72	0.46
	POA (%)	44	56	41	69	59	59	63	44	59	85	81	60
Rater 15	**Kappa**	0.72	0.63	0.27	0.78	0.64	0.77	0.78	0.49	0.69	0.74	0.95	0.68
	POA (%)	79	79	37	82	75	86	86	71	79	86	96	78
Rater 20	**Kappa**	0.75	0.66	0.52	0.73	0.70	0.59	0.49	0.75	0.64	0.51	0.88	0.66
	POA (%)	81	78	59	78	78	70	67	85	74	70	93	76
Rater 21	**Kappa**	0.41	0.39	0.27	0.49	0.50	0.39	0.53	0.22	0.41	0.31	0.72	0.42
	POA (%)	61	61	43	57	64	54	68	61	54	50	81	59
Rater 23	**Kappa**	0.52	0.52	0.44	0.64	0.70	0.56	0.72	0.51	0.64	0.63	0.88	0.61
	POA (%)	63	74	56	70	78	74	81	81	74	78	93	75
Rater 24	**Kappa**	0.59	0.53	0.40	0.62	0.79	0.84	0.94	0.66	0.88	1.00	0.94	0.75
	POA (%)	59	69	50	70	85	89	96	77	93	100	96	80
Rater 26	**Kappa**	0.52	0.49	0.39	0.48	0.72	0.49	0.45	0.56	0.69	0.71	0.83	0.58
	POA (%)	64	68	50	57	79	64	61	71	79	86	89	70
Rater 28	**Kappa**	0.58	0.50	0.30	0.69	0.73	0.54	0.58	0.72	0.58	0.59	0.83	0.60
	POA (%)	67	67	41	78	81	67	70	81	70	78	89	72
Mean per joint	**Kappa**	0.57	0.55	0.44	0.66	0.70	0.64	0.69	0.59	0.64	0.68	0.86	
	POA (%)	68	72	53	73	78	75	79	75	74	81	91	

FPA=foot progression angle; ST=during stance; SW=during swing; POA=percentage of agreement.

Table S8. Interrater agreement kappa and POA for each experienced rater vs. criterion classification (N=15).

		Sagittal plane						Coronal plane		Transverse plane			
		Pelvis	**Hip**	**Knee ST**	**Knee SW**	**Ankle ST**	**Ankle SW**	**Pelvis**	**Hip**	**Pelvis**	**Hip**	**FPA**	**Mean per rater**
Rater 1	**Kappa**	0.32	0.60	0.33	0.60	0.65	0.47	0.68	0.75	0.69	0.58	0.83	0.59
	POA (%)	44	78	44	67	74	63	78	89	78	74	89	71
Rater 2	**Kappa**	0.67	0.53	0.52	0.78	0.61	0.46	0.56	0.48	0.64	0.59	0.89	0.61
	POA (%)	75	71	61	82	71	64	68	68	75	75	93	73
Rater 3	**Kappa**	0.46	0.61	0.61	0.58	0.55	0.71	0.56	0.20	0.67	0.72	1.00	0.61
	POA (%)	59	74	67	67	67	81	70	48	77	85	100	72
Rater 5	**Kappa**	0.62	0.57	0.52	0.58	0.73	0.90	0.65	0.70	0.78	0.54	0.94	0.68
	POA (%)	70	74	59	67	81	93	74	81	85	74	96	78
Rater 6	**Kappa**	0.82	0.66	0.57	0.90	0.56	0.89	0.73	1.00	0.73	0.78	0.94	0.78
	POA (%)	89	78	63	93	67	93	81	100	81	89	96	85
Rater 8	**Kappa**	0.73	0.62	0.73	0.59	0.71	0.41	0.48	0.63	0.74	0.74	1.00	0.67
	POA (%)	81	78	78	67	78	67	70	78	81	89	100	79
Rater 10	**Kappa**	0.63	0.61	0.25	0.61	0.94	0.66	0.73	0.53	0.69	0.50	0.89	0.64
	POA (%)	70	74	37	70	96	74	81	70	78	74	93	74
Rater 16	**Kappa**	0.71	0.69	0.47	0.56	0.68	0.66	0.65	0.13	0.64	0.69	0.89	0.61
	POA (%)	79	82	57	64	79	79	79	79	75	82	93	77
Rater 17	**Kappa**	0.61	0.44	0.69	0.49	0.79	0.90	0.52	0.62	0.57	0.76	0.89	0.66
	POA (%)	70	63	74	59	85	93	67	78	70	89	93	76
Rater 18	**Kappa**	0.61	0.65	0.57	0.69	0.70	0.79	0.48	0.83	0.55	0.80	1.00	0.70
	POA (%)	74	77	63	74	78	85	62	89	67	89	100	78
Rater 19	**Kappa**	0.33	0.56	0.44	0.66	0.65	0.63	0.72	0.43	0.53	0.93	0.94	0.62
	POA (%)	44	74	52	74	74	74	81	67	67	96	96	73
Rater 22	**Kappa**	0.53	0.74	0.31	0.59	0.57	0.63	0.84	0.66	0.45	0.68	1.00	0.64
	POA (%)	67	85	41	67	67	78	89	85	59	81	100	74
Rater 25	**Kappa**	0.51	0.55	0.52	0.68	0.57	0.84	0.83	0.58	0.72	1.00	1.00	0.71
	POA (%)	70	74	59	70	67	89	89	74	81	100	100	79
Rater 27	**Kappa**	0.77	0.44	0.49	0.74	0.64	0.84	0.95	0.77	0.84	0.91	1.00	0.76
	POA (%)	82	68	57	79	75	89	96	86	89	96	100	83
Rater 29	**Kappa**	0.60	0.68	0.71	0.79	0.85	0.76	0.95	0.88	0.95	0.83	0.94	0.81
	POA (%)	75	82	75	82	89	86	96	93	96	93	96	88
Mean per joint	**Kappa**	0.59	0.60	0.51	0.66	0.68	0.70	0.69	0.61	0.68	0.74	0.94	0.67
	POA (%)	70	75	59	72	77	80	79	79	77	86	96	

FPA=foot progression angle; ST=during stance; SW=during swing; POA=percentage of agreement.

Chapter 5

Statistical parametric mapping to identify differences between consensus-based gait patterns in children with cerebral palsy

Angela Nieuwenhuys
Eirini Papageorgiou
Kaat Desloovere
Guy Molenaers
Tinne De Laet

Journal of Biomechanics [in preparation]

Abstract

This study aimed to provide objective evidence for the existence of joint gait patterns in cerebral palsy (CP), which were subjectively defined by an expert panel via a recently developed Delphi consensus study. To do so, statistical parametric mapping (SPM) was used to compare the mean kinematic waveforms of 154 trials of typically developing (TD) children to the mean kinematic waveforms of 1719 trials of children with spastic CP, which were classified according to the classification rules of the Delphi study. Two main hypotheses stated that:

(a) patterns with minor gait deviations (n=11) do not differ significantly from the gait pattern of TD children;
(b) all other pathological patterns (n=38) differ from TD gait and the locations of difference within the gait cycle that are highlighted by SPM, concur with the locations described in the classification rules.

This study provided objective evidence toward the content validity of the examined gait patterns in CP. Most patterns with 'no or minor gait deviations' (n=11) differed somewhat unexpectedly from TD gait, but these differences were generally small (≤3°). Further evidence demonstrated that the other pathological joint patterns (n=38) differed from TD gait and from each other. The locations within the gait cycle where gait patterns differed significantly from TD gait coincided well with the subjective, consensus-based classification rules. Nonetheless, some additional areas, which were not included within the pattern definitions of the consensus study, were also highlighted by the SPM analysis. Based on these results, suggestions to improve current pattern definitions were made.

Introduction

Three-dimensional gait analysis (3DGA) serves as a golden standard to objectively evaluate pathological gait in children with cerebral palsy (CP) and it has been shown to alter clinical decision making and improve treatment outcome[1–4]. However, the clinical interpretation of kinematic and kinetic gait data is subjective and therefore less reliable[5]. To support the clinical understanding of gait data, many attempts have been made to recognize different gait patterns from kinematic and kinetic reports, using either qualitative or quantitative approaches[6–13]. Regarding quantitative approaches, complex clinical interpretation of the patterns hinders the applicability of the classifications in medical practice. Qualitative approaches have also been criticized for unclear pattern definitions and lack of transparency in the construction of the classification[6]. In an attempt to overcome some of these methodological challenges, gait patterns for the different lower limb joints have recently been proposed for children with CP following an international consensus study[11]. Based on the judgment of an expert panel and supported by literature, three to seven nominal patterns were defined for the pelvis and hip joints across the three anatomical planes, for the ankle and knee during stance and swing phase in the sagittal plane, and for the foot progression angle. The pattern definitions or classification rules on which consensus was achieved were based on kinematic descriptions of locations within the gait cycle that deviate from the gait pattern of typically developing (TD) children. To a lesser extent, pattern definitions also included kinetic abnormalities for the hip patterns and for the knee patterns during stance in the sagittal plane.

The pattern definitions were the result of an informed, yet subjective opinion of an expert panel and therefore may provide an incomplete picture on gait patterns in CP. Hence, the content validity of the classification could be threatened and objective, quantitative data should be provided to support the identification of these consensus-based patterns. To this end, the present study uses statistical parametric mapping (SPM), a statistical approach which allows hypothesis testing on kinematic and kinetic waveforms without the need of a priori data reduction[14]. SPM is used to analyze kinematic and kinetic gait trials that were classified according to the definitions of the consensus study, in a large cohort of children with CP. If the classification has good content validity, all clinically relevant gait deviations should be included in the pattern definitions and all ambulatory children with spastic CP should be classifiable by a clinician, fitting the classification rules. Joint kinematics that do not fit any

pathological pattern, should therefore be classified as having 'no or minor gait deviations', which is a pattern that was defined at the level of each joint[11]. As a result, the **first hypothesis** stated that all kinematic trials classified as 'no or minor gait deviations' do not differ significantly from the gait pattern of TD children. The **second hypothesis** stated that kinematic and kinetic trials classified as pathological joint patterns, are significantly different from the gait of TD children and that the locations of difference within the gait cycle, which are highlighted by SPM, concur with the locations described in the classification rules of the consensus study[11]. A confirmation of this second hypothesis provides evidence for the feasibility of developing algorithms for automatic classification (e.g. Bayesian networks[13,15]) and for the classes and classification rules of the consensus study[11]. In light of this, **the third hypothesis** stated that the pathological patterns at the level of each joint differ from each other during at least one phase of the gait cycle.

Methodology

Patient group

The database of the clinical motion analysis laboratory of University Hospital Pellenberg was searched for gait analysis sessions of children with unilateral or bilateral spastic CP, aged between 3 to 18 years and GMFCS level I, II, or III. Children with marked signs of dystonia or ataxia were excluded, but any previous treatments were allowed. To compare pathological gait to the normal gait pattern, the reference database of the hospital was used, which consisted of 56 TD children between 5 to 18 years old, with no history of musculoskeletal or neuromotor disorders.

Data collection

Standardized 3DGA measurements were performed using ten to fifteen optoelectronic cameras (Vicon Motion Systems, Oxford, UK) and two force platforms (Advanced Mechanical Technology Inc., USA), which were embedded in a 10m walkway. Reflective markers were fixed on anatomical landmarks according to the Plug-In-Gait model and all children were asked to walk barefoot and at a self-selected speed. Nexus software was used to estimate gait cycles, joint angles, and joint moments, which were normalized to body mass. Kinematic and kinetic waveforms were also time-normalized to the gait cycle, or to stance and swing phase when appropriate. Each waveform was interpolated to intervals of 2%,

yielding a total of 51 data points per curve. Subsequently, these kinematic and kinetic trials were imported into a custom-made Matlab® software tool. Trials with artifacts or with signs of inaccurate marker placement were excluded, as well as trials that were not representative of a child's gait. For each TD child, two to four good quality trials of the left or right side were included for SPM analysis. For patients with CP, all available trials were included in the study. A median of 3 trials (interquartile range 2 tot 7) were available for classification per patient per side. The maximum number of trials for each patient and TD child was included because the research question of this methodological study concerns the analysis of differences between kinematic and kinetic groups as they are defined subjectively by clinicians, irrespective of whether trials belong to the same patient or different patients, unlike for an analysis of the prevalence of the gait patterns within different patient groups, which would require a fixed number of trials (or an averaged trial) per patient. Following the definitions of the consensus study[11], each available trial for each included CP patient was classified by a clinical expert rater (one of two raters) for the following joints: pelvis in the sagittal (PS), coronal (PC), and transverse (PT) plane; hip in the sagittal (HS), coronal (HC), and transverse (HT) plane; knee during stance (KSTS) and during swing (KSWS) in the sagittal plane; ankle during stance (ASTS) and during swing (ASWS) in the sagittal plane, and foot progression angle (FPA). A brief description of each pattern (n=49) is presented in Table 1.

Statistical analysis

To test the first and second hypothesis, SPM unpaired t-tests were performed, comparing the mean kinematic (or kinetic) angle of each pattern to the respective mean kinematic (or kinetic) angle of the TD group (α=0.01). For the third hypothesis, an SPM one-way-ANOVA was performed to examine whether the mean joint angles of the patterns per joint differed significantly from each other (α=0.01).

For each SPM ANOVA or t-test, a statistical parametric map (SPM{F} or SPM{t} respectively) was created by calculating the conventional univariate t- or F-statistic at each point of the gait curve[14]. Afterwards, Random Field Theory allowed an estimation of the critical threshold above which only 1% (α=0.01) of equally smooth random data was expected to cross[16]. If the SPM{F} crossed the critical threshold, post-hoc SPM{t} maps were calculated for between-group comparisons. If at any time, an SPM{t} crossed the critical threshold, a supra-threshold cluster was created, indicating a significant difference between

two gait patterns in a specific location of the gait cycle. A Bonferroni correction was applied for each joint to adjust α for multiple post-hoc comparisons. For each supra-threshold cluster, the probability (p-value) of discovering a cluster with similar proportions when testing equally smooth random data was calculated[16]. Because of the high number of statistical analyses, the SPM results are presented in a summarized manner. Instead of SPM{t} curves, black bars will be shown, indicating the locations within the gait cycle during which a supra-threshold cluster was identified (Figure 1). Taking into account previously reported measurement errors that are inherent to 3DGA, a significant difference was interpreted as relevant if the mean waveforms were at least 3° removed from each other within the areas of significance as indicated by the SPM output (i.e. black bars)[17,18]. All analyses were performed using open-source SPM1d code (vM.01.0003; www.spm1D.org) in Matlab® and the study was approved by the Medical Ethical Committee of University Hospitals Leuven (s56036).

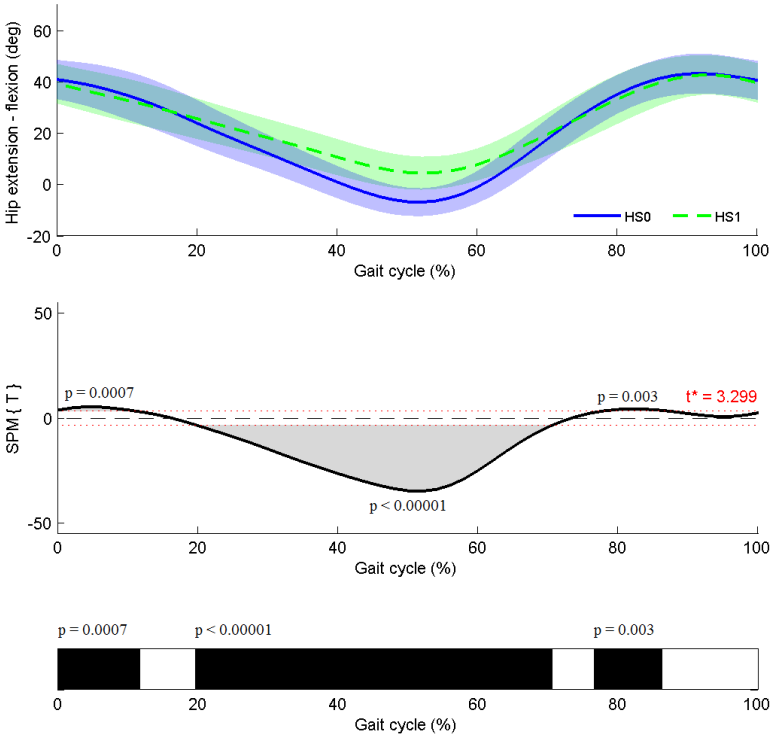


Figure 1. Upper graph shows the mean kinematic hip angle in the sagittal plane of trials classified as HS0 (no or minor gait deviations) or HS1 (hip extension deficit). Middle graph shows SPM {t} statistic as a function of the gait cycle. The critical threshold (t*) was exceeded between 0-12%, 20-71%, and 76-86% of the gait cycle. Lower black bars represent a simplified visualization of the significant areas indicated by the SPM{t} statistic.

Table 1. Observed frequency (%) and brief description of all sagittal, coronal, and transverse plane joint patterns defined during the consensus study.

SAGITTAL PLANE	**(%)**
Pelvis	
PS0 - Normal pelvic motion/posture – no or minor gait deviations	16.0
PS1 - Increased range of motion	29.6
PS2 - Increased anterior tilt on average	16.1
PS3 - Increased anterior tilt and increased range of motion	35.9
PS4 - Decreased anterior tilt (posterior tilt)	1.3
PS5 - Decreased anterior tilt (posterior tilt) and increased range of motion	1.1
Hip	
HS0 - Normal hip motion – no or minor gait deviations	55.3
HS1 - Hip extension deficit	27.6
HS2 - Continuous excessive hip flexion	17.1
Knee during stance	
KSTS0 - Normal knee motion during stance – no or minor gait deviations	14.8
KSTS1 - Increased knee flexion at initial contact	7.3
KSTS2 - Increased knee flexion at initial contact and earlier knee extension movement	20.7
KSTS3 - Knee hyperextension	8.1
KSTS4 - Knee hyperextension and increased knee flexion at initial contact	10.9
KSTS5 - Increased flexion in midstance and internal flexion moment present	23.0
KSTS6 - Increased flexion in midstance and internal extension moment present	15.1
Knee during swing	
KSWS0 - Normal knee motion during swing – no or minor gait deviations	35.4
KSWS1 - Delayed peak knee flexion	21.5
KSWS2 - Increased peak knee flexion	12.6
KSWS3 - Increased and delayed peak knee flexion	9.4
KSWS4 - Decreased peak knee flexion	10.8
KSWS5 - Decreased and delayed peak knee flexion	10.3
Ankle during stance	
ASTS0 - Normal ankle motion during stance – no or minor gait deviations	38.6
ASTS1 - Horizontal second ankle rocker	28.0
ASTS2 - Reversed second ankle rocker	9.4
ASTS3 - Equinus gait	4.2
ASTS4 - Calcaneus gait	19.7
Ankle during swing	
ASWS0 - Normal ankle motion during swing – no or minor gait deviations	40.0
ASWS1 - Insufficient prepositioning in terminal swing	6.5
ASWS2 - Continuous plantarflexion during swing (drop foot)	18.7
ASWS3 - Excessive dorsiflexion during swing	34.8

Table 1. Continued.

CORONAL PLANE	**(%)**
Pelvis	
PC0 - Normal pelvic motion/posture – no or minor gait deviations	48.6
PC1 - Increased pelvic range of motion	29.1
PC2 - Continuous pelvic elevation	11.8
PC3 - Continuous pelvic depression	10.6
Hip	
HC0 - Normal hip motion – no or minor gait deviations	62.9
HC1 - Excessive hip abduction in swing	21.6
HC2 - Continuous excessive hip abduction	9.2
HC3 - Continuous excessive hip adduction	6.3
TRANSVERSE PLANE	
Pelvis	
PT0 - Normal pelvic motion/posture – no or minor gait deviations	44.4
PT1 - Increased pelvic range of motion	30.4
PT2 - Excessive pelvic external rotation during the gait cycle	13.0
PT3 - Excessive pelvic internal rotation during the gait cycle	12.2
Hip	
HT0 - Normal hip motion – no or minor gait deviations	75.4
HT1 - Excessive hip external rotation during the gait cycle	8.9
HT2 - Excessive hip internal rotation during the gait cycle	15.7
Foot	
FPA0 - Normal foot progression angle – no or minor gait deviations	66.6
FPA1 - Outtoeing	15.7
FPA2 - Intoeing	17.7

Described deviations such as increased or excessive joint angles refer to deviations which are more than one standard deviation away from the TD reference database. A more detailed description of the patterns is available in Nieuwenhuys et al.[11].

Results

Table 2 describes the characteristics of CP and TD children. In total, 459 gait analysis sessions corresponding to 356 CP patients were included of which 154 sessions were post-treatment (i.e. Botulinum toxin type A injection session, selective dorsal rhizotomy, or single event multilevel surgery). One gait analysis session was available for 275 patients; two sessions were available for 67 patients, and three to six sessions for 14 patients. Overall, 1719 good quality kinematic trials were classified of which 985 also had kinetic data available.

Regarding the TD children, 154 good quality kinematic trials were included for SPM analysis, of which 148 trials also included the kinetic data.

Table 2. Demographic characteristics of CP (n=356) and TD (n=56) group.

	CP (n)	**TD (n)**
Gender		
Male	212	24
Female	144	32
Weight (mean (SD), in kg)	32.2 (14.0)	40.1 (17.7)
Height (mean (SD), in m)	1.34 (0.20)	1.48 (0.21)
Diagnosis		
Bilateral CP	219	
Unilateral CP	137	
GMFCS		
Level I	192	
Level II	117	
Level III	47	
Number of 3DGA sessions	459	56
Age at time of 3DGA (mean (SD), in years)	9 years, 10 months (3 years, 6 months)	11 years, 1 month (3 years, 10 months)

SD = standard deviation.

Hypothesis 1: kinematic trials classified as 'no or minor gait deviations' at the level of each joint do not differ significantly from the gait pattern of TD children

For all except one of the joints, the pattern with 'no or minor gait deviations' differed significantly from TD gait during at least one phase of the gait cycle (all $p<0.01$; Figure 2). Only for the pattern of the pelvis in the transverse plane, no significant differences were identified. In the locations of the gait cycle where significant deviations from TD gait were identified, the differences between the mean kinematic angles were generally small (≤3°). Larger deviations from TD gait were identified for the pattern with 'no or minor gait deviations' of HS, KSWS, ASTS, and ASWS. Compared to TD gait, increased hip flexion was noted between 0-61% and 74-100% of the gait cycle (both $p<0.00001$) and the knee flexion angle during swing was also increased between 0-24% and 49-100% of swing ($p=0.00227$ and $p<0.00001$ respectively). At the level of the ankle, the pattern "no or minor gait deviations" showed markedly increased dorsiflexion during push-off (82-100% of stance,

p=0.0013) and during the first 29% of swing (p<0.0001). A slight increase in dorsiflexion was also identified between 4-29% of stance phase (p<0.0001).

Hypothesis 2: kinematic and kinetic trials classified as one of the pathological joint patterns are significantly different from TD gait; locations of difference within the gait cycle that are highlighted by SPM concur with the locations described in the classification rules of the consensus study

All pathological patterns differed significantly from TD gait, on average throughout 91% of the gait cycle (or of stance/swing phase regarding the patterns of FPA, KSTS, KSWS, ASTS, and ASWS). Locations of difference that were highlighted by SPM concurred with the locations described in the classification rules for all patterns of the following joints: FPA, HT, PS, PC, and PT (Figures S1-S5). Overall, mostly large significant differences (>3°) between the mean TD pattern and the pathological patterns of these joints were identified. SPM analysis only highlighted small (≤3°) differences from TD gait for PC1 and PT1, indicating 'increased pelvic range of motion' in the coronal and transverse plane.

As regards the other joints, statistical analyses identified at least one pattern per joint that concurred with the classification rules, but also yielded additional significant locations which were large (>3°) and were not incorporated in the pattern definitions. Firstly, HS1 or 'hip extension deficit', which is defined based on stance phase kinematic deviations, also presented excessive hip flexion during 80-100% of swing phase (p=0.00004; Figure 3). Secondly, the knee patterns during swing are defined based on an abnormal peak flexion angle. In addition to this feature, these patterns presented with insufficient knee extension during the second half of swing phase (all p<0.00001; Figure 4). Thirdly, the definition of KSTS5 and KSTS6 is identical in terms of kinematic deviations (i.e. 'increased knee flexion in midstance'). On top of excessive knee flexion in midstance, both patterns were observed with significantly increased knee flexion compared to TD gait over the entire stance phase (both p<0.00001; Figure 5). Subsequently, KSTS1, which was defined as 'increased knee flexion at initial contact', was further observed to have significantly increased knee flexion between 0-71% of stance phase (p<0.00001). KSTS2 was defined as 'increased flexion at initial contact and earlier knee extension movement' and additionally showed significantly increased knee flexion between 53-92% of stance (p=0.00001). However, the difference between KSTS2 and TD gait during this phase was small (≤3°). Fourthly, the ankle patterns representing a 'horizontal' or 'reversed second ankle rocker' (ASTS1 and ASTS2)

additionally presented with significantly increased dorsiflexion during loading response compared to TD gait (both $p<0.001$; Figure 6). Furthermore, ASTS1, ASTS2, and ASTS4 differed from TD gait during pre-swing (all $p<0.01$). Fifthly, it was observed that ASWS1, which is defined as 'insufficient preposition in terminal swing', also showed insufficient plantarflexion between 0-27% of swing ($p=0.00083$; Figure 7). Sixthly, in the coronal plane, HC1, defined as 'excessive hip abduction during swing', further showed excessive abduction between 0-35% of the gait cycle ($p<0.00001$) and slightly increased (≤3°) adduction between 49-67% of the gait cycle ($p=0.00031$; Figure 8) compared to TD gait.

The mean kinetic curves of the patterns that contain a description of kinetic deviations (HS1, KSTS 3-4-5-6) were all found to differ significantly from their respective TD joint moments (Figure S6-7). The locations of difference concurred with the classification rules. In addition, small (≤3°) significant locations were identified for each of those patterns during the first 15% of stance phase.

Hypothesis 3: the kinematic and kinetic trials of the gait patterns at the level of each joint are different from each other in at least one part of the gait cycle.

SPM ANOVAs of kinematic and kinetic trials identified significant differences between the patterns of each joint ($p<0.01$). Post-hoc SPM t-tests indicated that 'decreased pelvic anterior tilt' (PS4) and 'decreased pelvic anterior tilt and increased range of motion' (PS5) did not differ significantly from each other throughout the gait cycle (Figure S3). The other patterns at the level of each joint were found to be significantly different from the other joint patterns, on average throughout 91 % of the gait cycle (or of stance/swing phase; Figures 3-8, S1-S7).

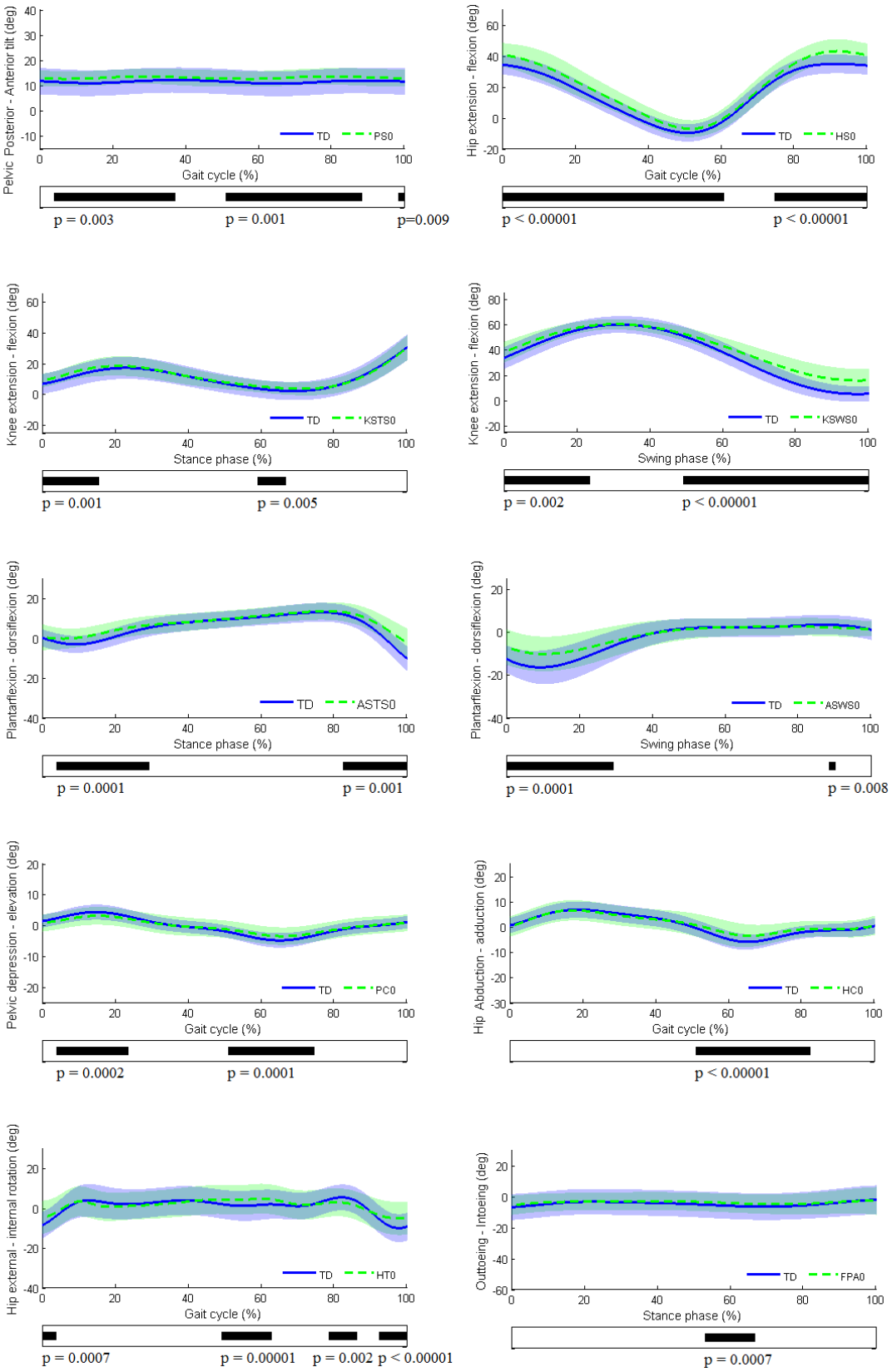


Figure 2. Graphs show the mean kinematic angle of TD gait versus the pattern 'no or minor gait deviations' at the level of each joint, except for PT (no significant differences). Black bars indicate gait phases during which the SPM{t} statistic exceeded the critical threshold.

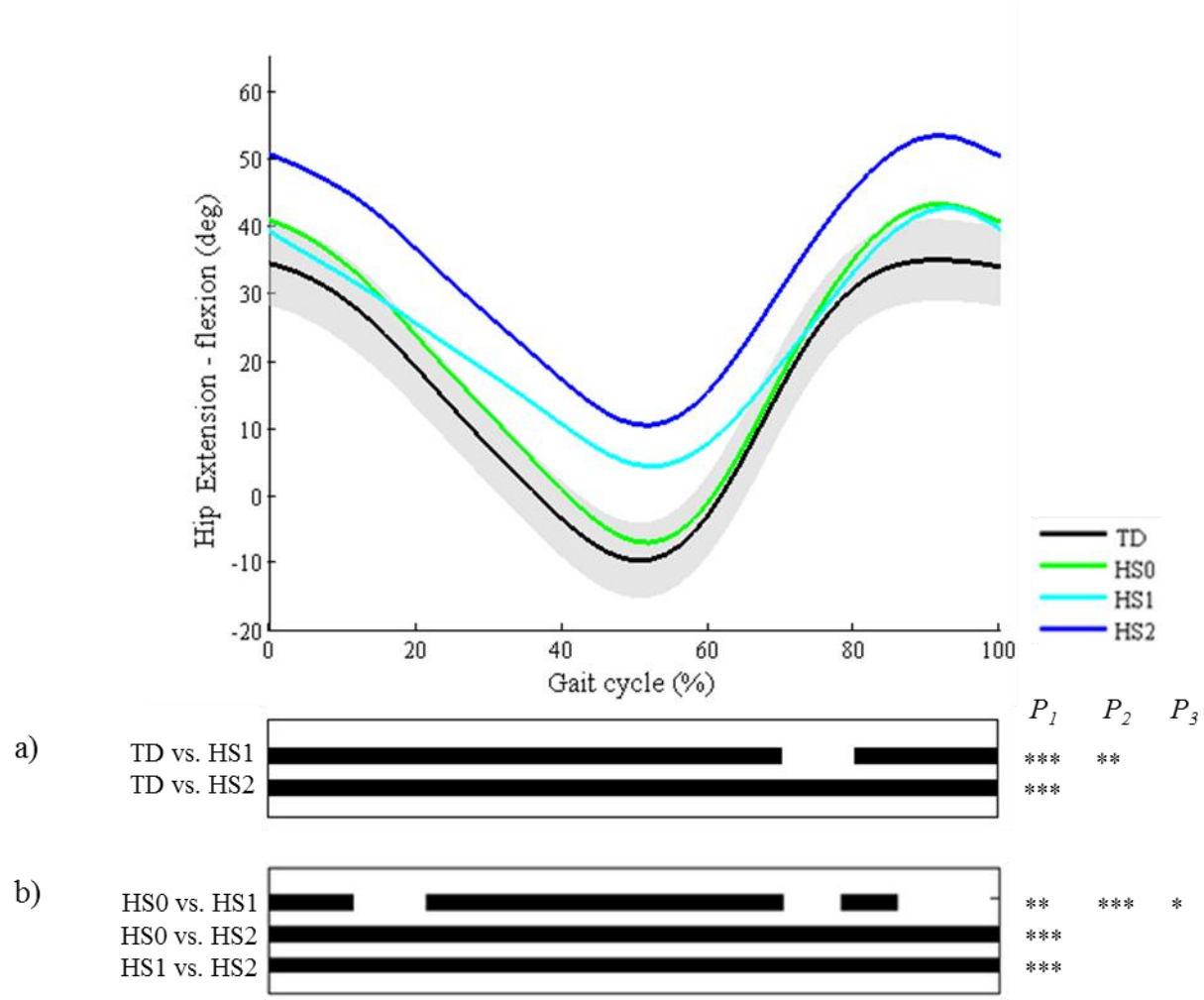


Figure 3. Top graph shows the mean kinematic angle of TD gait and of each consensus-based pattern at the level of the hip in the sagittal plane (HS). Black bars indicate significant gait phases during which the SPM{t} statistic exceeded the critical threshold. Panel (**a**) shows results of hypothesis 2 (i.e. unpaired t-tests, α=0.01); panel (**b**) shows results of hypothesis 3 (i.e. post-hoc unpaired t-tests, α=0.003). * $p<0.01$, ** $p<0.001$, *** $p<0.00001$. P_1 indicates the p-value of the first cluster during the gait cycle, P_2 the second cluster, etc.

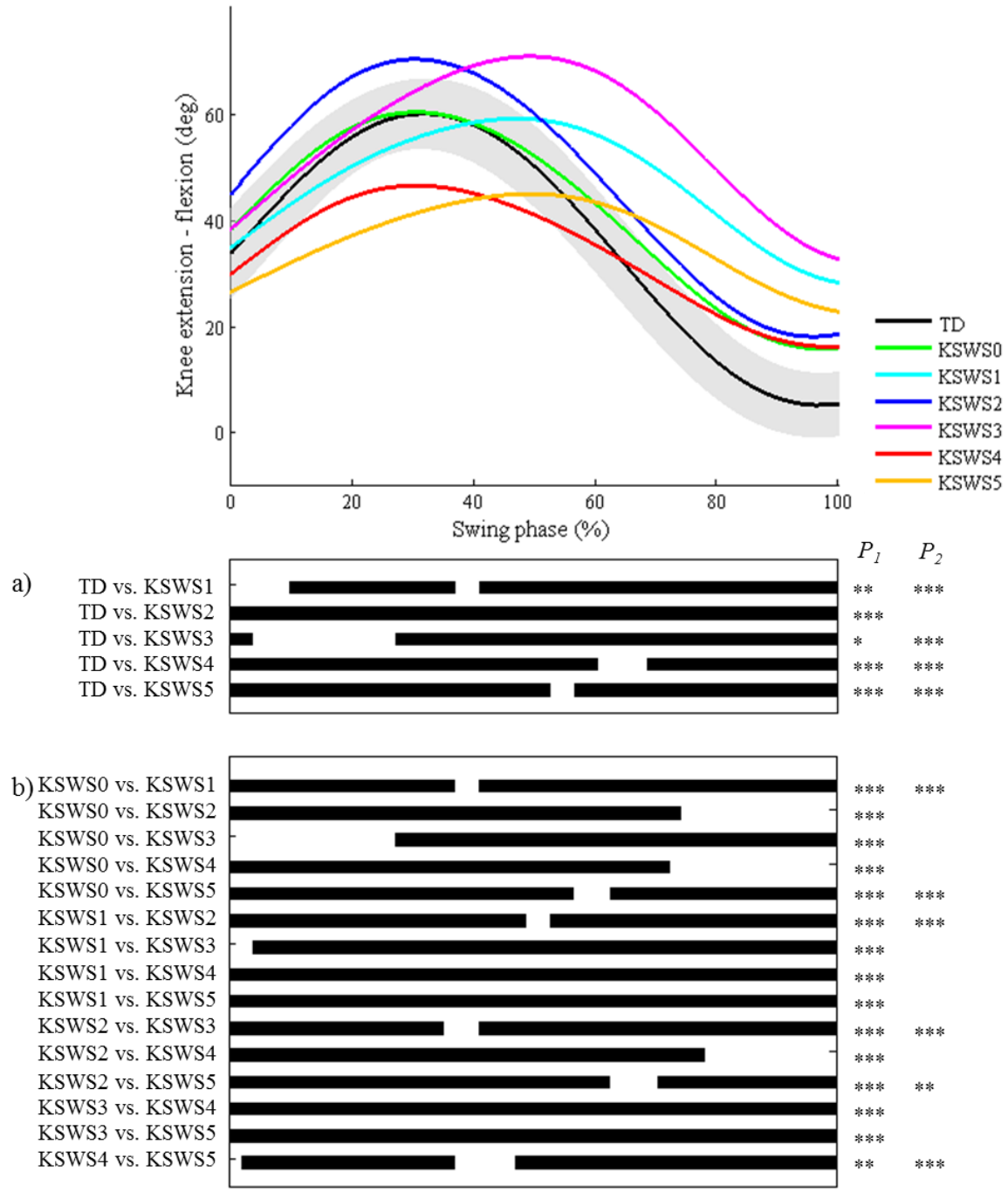


Figure 4. Top graph shows the mean kinematic angle of TD gait and of each consensus-based pattern at the level of the knee during swing phase in the sagittal plane (KSWS). Black bars indicate significant gait phases during which the SPM{t} statistic exceeded the critical threshold. Panel (**a**) shows results of hypothesis 2 (i.e. unpaired t-tests, α=0.01); panel (**b**) shows results of hypothesis 3 (i.e. post-hoc unpaired t-tests, α=0.0006). * $p<0.01$, ** $p<0.001$, *** $p<0.00001$. P_1 indicates the p-value of the first cluster during the swing phase, P_2 the second cluster.

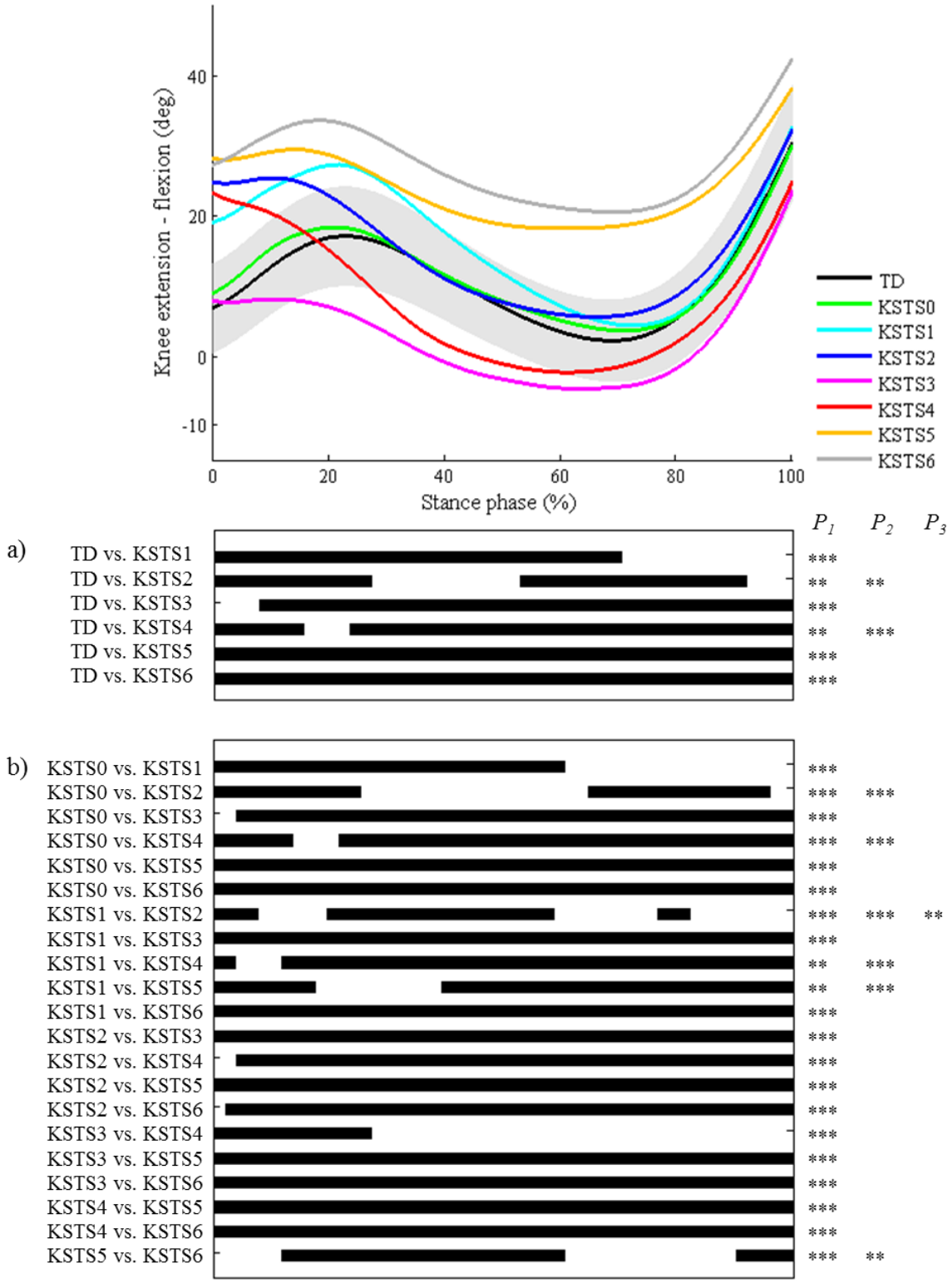


Figure 5. Top graph shows the mean kinematic angle of TD gait and of each consensus-based pattern at the level of the knee during stance phase in the sagittal plane (KSTS). Black bars indicate significant gait phases during which the SPM{t} statistic exceeded the critical threshold. Panel (**a**) shows results of hypothesis 2 (i.e. unpaired t-tests, α=0.01); panel (**b**) shows results of hypothesis 3 (i.e. post-hoc unpaired t-tests, α=0.0005). * $p<0.01$, ** $p<0.001$, *** $p<0.00001$. P_1 indicates the p-value of the first cluster during the stance phase, P_2 the second cluster, etc.

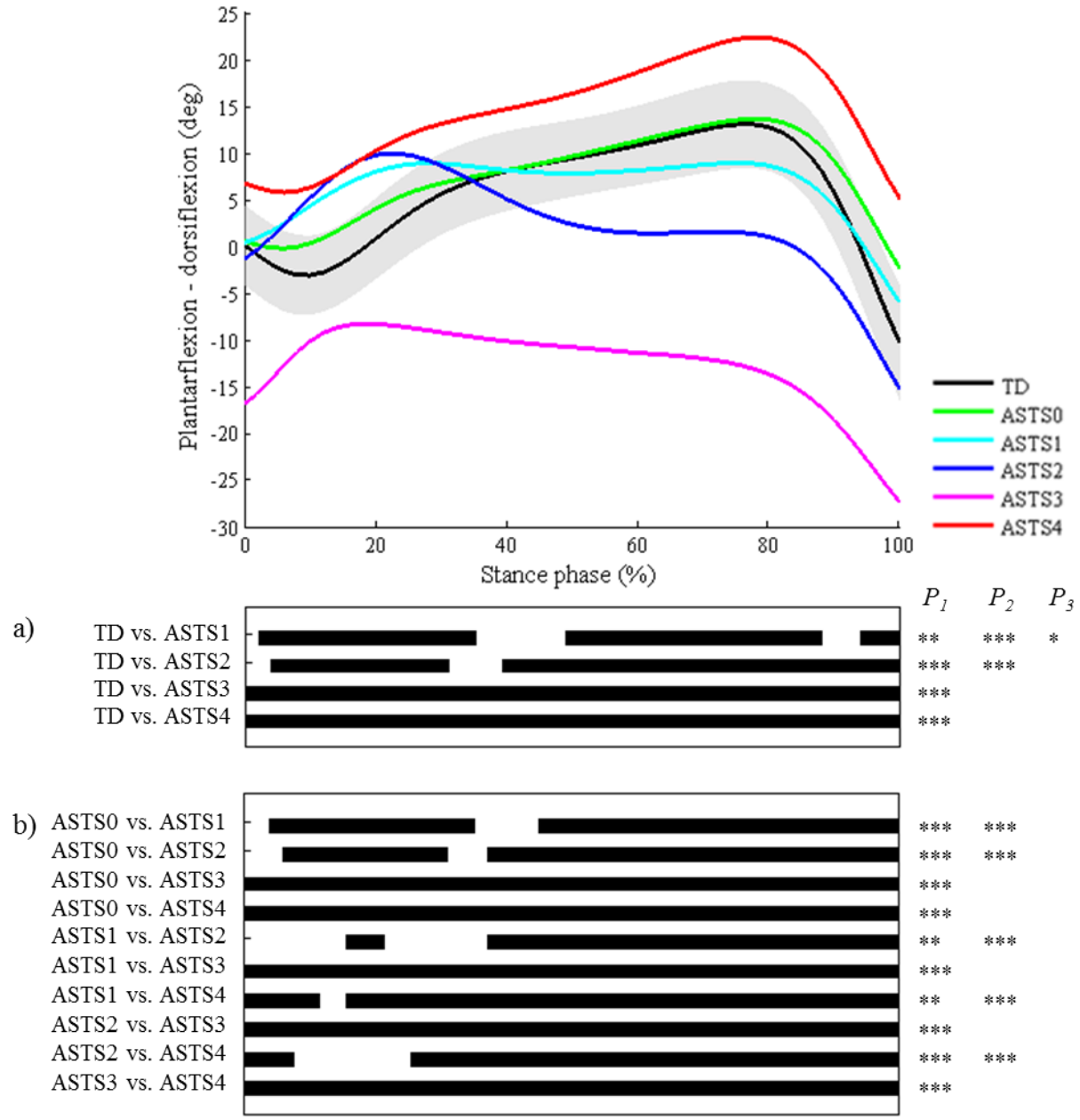


Figure 6. Top graph shows the mean kinematic angle of TD gait and of each consensus-based pattern at the level of the ankle during stance phase in the sagittal plane (ASTS). Black bars indicate significant gait phases during which the SPM{t} statistic exceeded the critical threshold. Panel (**a**) shows results of hypothesis 2 (i.e. unpaired t-tests, α=0.01); panel (**b**) shows results of hypothesis 3 (i.e. post-hoc unpaired t-tests, α=0.001). * p<0.01, ** p<0.001, *** p<0.00001. P_1 indicates the p-value of the first cluster during the stance phase, P_2 the second cluster, etc.

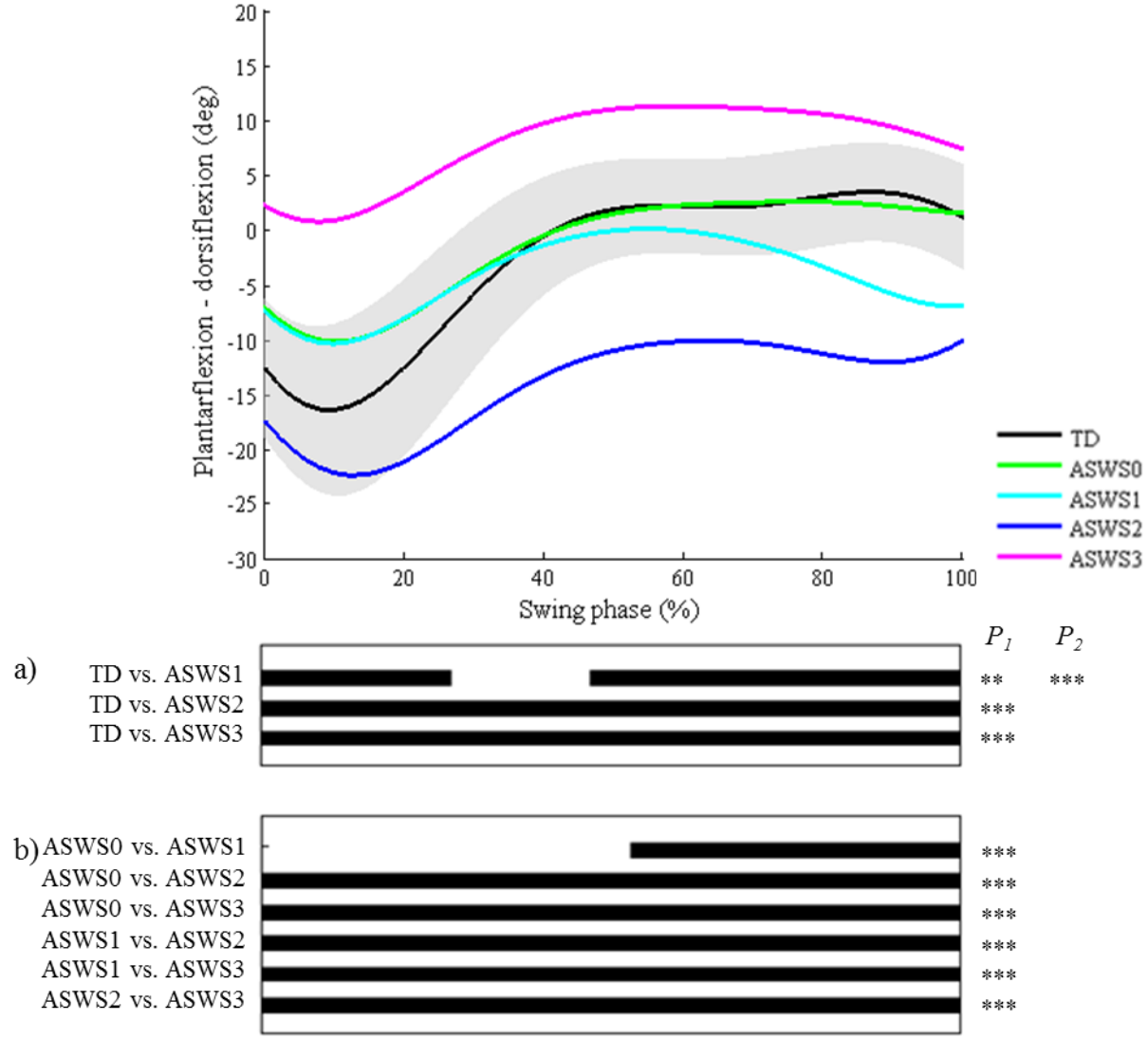


Figure 7. Top graph shows the mean kinematic angle of TD gait and of each consensus-based pattern at the level of the ankle during swing phase in the sagittal plane (ASWS). Black bars indicate significant gait phases during which the SPM{t} statistic exceeded the critical threshold. Panel (**a**) shows results of hypothesis 2 (i.e. unpaired t-tests, α=0.01); panel (**b**) shows results of hypothesis 3 (i.e. post-hoc unpaired t-tests, α=0.002). * p<0.01, ** p<0.001, *** p<0.00001. P_1 indicates the p-value of the first cluster during the swing phase, P_2 the second cluster.

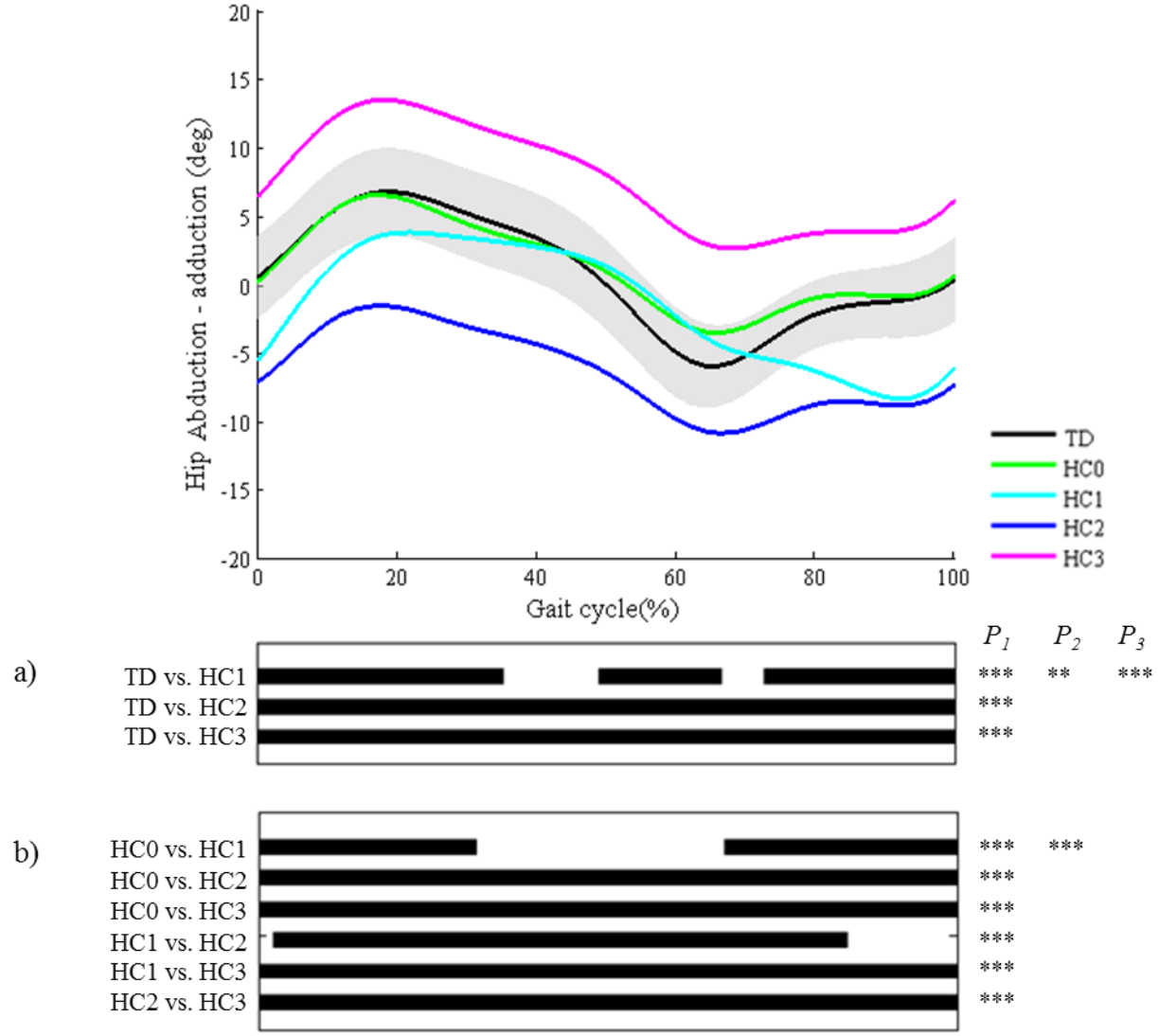


Figure 8. Top graph shows the mean kinematic angle of TD gait and of each consensus-based pattern at the level of the hip in the coronal plane (HC). Black bars indicate significant gait phases during which the SPM{t} statistic exceeded the critical threshold. Panel (**a**) shows results of hypothesis 2 (i.e. unpaired t-tests, α=0.01); panel (**b**) shows results of hypothesis 3 (i.e. post-hoc unpaired t-tests, α=0.002). * $p<0.01$, ** $p<0.001$, *** $p<0.00001$. P_1 indicates the p-value of the first cluster during the gait cycle, P_2 the second cluster, etc.

Discussion

This study examined the content validity of a recently published gait classification for children with spastic CP. The purpose was to provide objective evidence for the existence of joint gait patterns in CP, which were developed and subjectively defined by an expert panel via a consensus study[11]. SPM was used to analyze a large database of kinematic and kinetic trials that were classified by clinicians to investigate three hypotheses.

The first hypothesis assumed that the patterns with 'no or minor gait deviations' at the level of each joint, would not differ from the gait pattern of TD children. This hypothesis could only be confirmed for the pelvis in the transverse plane. Since the pattern with minor gait deviations differed from TD gait for all other joints, it could be assumed that common gait deviations in CP were not included in the classification, which would threaten its content validity. However, for most joints, the deviations from the mean angle of TD gait were less than 3°. It can therefore be assumed that these differences are clinically of less relevance, especially when also taking into account possible inter-therapist or inter-session measurement errors[18]. On the other hand, the results also indicated significant areas during the gait cycle where the differences between TD gait and the pattern with minor gait deviations of the hip in the sagittal plane, of the knee during swing, and of the ankle during stance and swing were more meaningful (significant and more than 3°). This could indicate that relevant information is not included in the pattern definitions or a potential pattern might have been missed. This is probably not the case for the observed increased hip flexion, because this deviation is incorporated in the patterns HS1 and HS2 and a patient will be classified as such if hip flexion would further increase. Similarly, regarding insufficient knee extension during terminal swing, one could argue that this important clinical information is already sufficiently represented in the knee patterns during stance that have the feature 'increased knee flexion at initial contact' (KSTS1, 2, and 4). However, this is not the case for the significant differences in the ankle patterns, which occurred during the first and third ankle rocker, as well as during early swing. Deviations in these locations of the gait cycle also appeared in the results for the second and third hypotheses as being discriminatory between different gait patterns. Specific kinematic deviations related to the first and third ankle rockers are currently not included in the pattern definitions of the studied classification, nor were they included in previously reported classifications[6–10]. It should be further investigated to what extent these locations can help improve the current patterns definitions.

The second hypothesis assumed that all other pathological patterns differed significantly from the gait pattern of TD children in the key locations of the gait cycle that were indicated in the pattern definitions by the experts. A general conclusion from the results is that for each pattern, all key locations that were originally included in the classification rules were indeed highlighted as significant areas by the SPM analysis. However, on several occasions, additional significant locations were indicated by SPM analysis during which patterns also differed from TD gait. This information could be used to further refine some pattern definitions. For instance, all knee patterns during swing were characterized by insufficient knee extension during terminal swing compared to TD gait. The results related to the third hypothesis (Figure 4) clearly highlighted that all patterns without the feature 'delayed peak knee flexion' (KSWS0-2-4) reached a similar knee flexion angle during terminal swing, which was significantly lower than the angles of KSWS1-3-5, but also approximately 10° higher than the angle of TD gait. If there is doubt about whether or not the peak knee flexion during swing is delayed, the knee angle during terminal stance could support the final choice. Regarding the knee pattern during stance, it was clear that patients, who fulfill the current criteria of excessive flexion during midstance, will also show excessive knee flexion during the remainder of stance (Figure 5). The kinematic deviations of KSTS5 and KSTS6 might therefore be redefined as 'continuously excessive knee flexion during stance', similar to the crouch pattern that was defined by Sutherland et al.[19]. The results related to the third hypothesis (Figure 5) indicated that KSTS6 also showed significantly higher knee flexion than KSTS5 between 10-67% of the stance phase, even though the definitions of these patterns in terms of kinematic deviations were identical in the consensus study[11]. The mean angle of KSTS6 reaches over 30° of knee flexion whereas the mean angle of KSTS5 does not. This information could help clinicians distinguish between KSTS5 and KSTS6 for patients that do not have kinetic data or trunk kinematics available, as trunk position will likely be an important factor influencing the generated knee moment during stance.

The **third hypothesis** assumed that all pathological patterns at the level of each joint are different from each other in at least one part of the gait cycle. Apart from two pelvic patterns in the sagittal plane, PS4 and PS5, this hypothesis was confirmed. The low observed frequency of these patterns (1.3% and 1.1% respectively) in this study might have limited the power of the SPM analysis to detect significant differences between both patterns. Also in literature, decreased pelvic tilt was not often described in CP gait classifications. The usefulness of these two patterns in the classification should therefore be questioned. Only

Rodda et al.[20] have mentioned decreased tilt as a possible feature of the Type IV gait pattern, which represents patients with severe crouch gait (i.e. excessive hip and knee flexion as well as excessive ankle dorsiflexion).

Regarding the statistical analyses, SPM unpaired t-tests were used for the first two hypotheses and SPM one-way-ANOVA was used to test the third. Alternatively, an SPM one-way-ANOVA could have been performed for each joint, including both the CP gait patterns and the TD gait trials. The post-hoc SPM t-tests would essentially constitute all comparisons that are reported in the present study, except that the critical threshold would be calculated based on a lower α-level because of the Bonferroni correction. To test whether this choice of statistics would have affected the conclusions, these analyses were also performed. Results showed that probabilities were lower and for several between-group comparisons the width of the clusters was slightly more narrow (generally for 2-4% of the gait cycle), but never to the extent that it would change the interpretation of the results. A limitation of this study is that the assumption of equal variance between all pathological patterns and TD gait could have been violated. It is possible that slightly higher critical thresholds would have been identified if corrections for unequal variances would have been performed, but this feature is challenging to be defined and was not available using the current SPM code for Matlab. Slightly stricter critical thresholds are not likely to alter the general conclusions of this study (cfr. supra), as the probability of most critical thresholds was very low ($p<0.00001$). A possible effect could be that some differences between TD gait and the patterns 'no or minor gait deviations' of FPA, ASWS, and KSTS might have been undetected, as the mean angles between these patterns and TD gait were smaller than 3° and probabilities for the supra-threshold clusters of these analyses were relatively close to 0.01 (Figure 2). Although all trials were considered independently, a potential learning effect could not be excluded as raters could not be blinded to patient identification. However, previous repeatability analyses suggested that this most likely did not influence the results[21].

In conclusion, the present study confirmed the content validity of the examined gait patterns in CP. It was found that most patterns with 'no or minor gait deviations' differed somewhat unexpectedly from TD gait, but differences were generally small (<3°). Further evidence demonstrated that the other pathological joint patterns differed from TD gait and from each other. The locations of significant difference between the patterns and TD gait coincided well with the subjective, consensus-based classification rules. Nonetheless, some additional areas,

which were not included within the pattern definitions of the consensus study, were also highlighted by the SPM analysis. Based on these results, suggestions to improve current pattern definitions were made. The results further suggest that algorithms, which could automate this classification[13], are likely to be successful. In a next step, it should be investigated to what extent the patterns are responsive to treatment and how they could be incorporated in the clinical reasoning process.

Conflict of Interest Statement

The authors declare to have no financial or personal relationships with other people or institutions that could be perceived as posing conflict or bias.

Acknowledgements

A. Nieuwenhuys is supported by an OT project of KU Leuven University (OT/12/100). E. Papageorgiou is supported by the MD Paedigree project, a Model-Driven Paediatric European Digital Repository, partially funded by the European Commission under FP7 – ICT Programme (grant agreement no: 600932, http://www.md-paedigree.eu) and by the SIMCP IWT-project, a simulation platform to predict gait performance following orthopedic intervention in children with cerebral palsy (IWT 140184). The funders had no role in the design of the study, nor in the collection, analysis, and interpretation of the study data. In addition, funders had no part in the writing of the manuscript or in the decision to submit the manuscript for publication.

References

1. Cook RE, Schneider I, Hazlewood ME, Hillman SJ, Robb JE. Gait Analysis Alters Decision-Making in Cerebral Palsy. *J Pediatr Orthop*. 2003:292-295.
2. De Luca P, Davis RB, Ounpuu S, Rose S, Sirkin R. Alterations in Surgical Decision Making in Patients with Cerebral Palsy based on Three-Dimensional Gait Analysis. *J Pediatr Orthop*. 1997;17(5):608-614.
3. Wren TAL, Gorton III GE, Ounpuu S, Tucker CA. Efficacy of clinical gait analysis: A systematic review. *Gait Posture*. 2011;34(2):149-153.
4. Wren TAL, Otsuka NY, Bowen RE, et al. Outcomes of lower extremity orthopedic surgery in ambulatory children with cerebral palsy with and without gait analysis: Results of a randomized controlled trial. *Gait Posture*. 2013;38(2):236-241.
5. Skaggs DL, Rethlefsen S a, Kay RM, Dennis SW, Reynolds R a, Tolo VT. Variability in gait analysis interpretation. *J Pediatr Orthop*. 2000;20(6):759-764.

6. Dobson F, Morris ME, Baker R, Graham HK. Gait classification in children with cerebral palsy: A systematic review. *Gait Posture*. 2007;25(1):140-152.

7. Carriero A, Zavatsky A, Stebbins J, Theologis T, Shefelbine SJ. Determination of gait patterns in children with spastic diplegic cerebral palsy using principal components. *Gait Posture*. 2009;29(1):71-75.

8. Toro B, Nester CJ, Farren PC. Cluster analysis for the extraction of sagittal gait patterns in children with cerebral palsy. *Gait Posture*. 2007;25(2):157-165.

9. Bonnefoy-Mazure A, Sagawa YJ, Lascombes P, De Coulon G, Armand S. Identification of gait patterns in individuals with cerebral palsy using multiple correspondence analysis. *Res Dev Disabil*. 2013;34(9):2684-2693.

10. Davids JR, Bagley AM. Identification of Common Gait Disruption Patterns in Children With Cerebral Palsy. *J Am Acad Orthop Surg*. 2014;22(12):782-790.

11. Nieuwenhuys A, Õunpuu S, Van Campenhout A, et al. Identification of joint patterns during gait in children with cerebral palsy: A Delphi consensus study. *Dev Med Child Neurol*. 2016;58(3):306-313.

12. Ferrari A, Alboresi S, Muzzini S, Pascale R, Perazza S, Cioni G. The term diplegia should be enhanced. Part I: a new rehabilitation oriented classification of cerebral palsy. *Eur J Phys Rehabil Med*. 2008;44(2):195-201.

13. Van Gestel L, De Laet T, Di Lello E, et al. Probabilistic gait classification in children with cerebral palsy: A Bayesian approach. *Res Dev Disabil*. 2011;32(6):2542-2552.

14. Pataky TC. Generalized n-dimensional biomechanical field analysis using statistical parametric mapping. *J Biomech*. 2010;43(10):1976-1982.

15. Zhang B, Zhang Y, Begg RK. Gait classification in children with cerebral palsy by Bayesian approach. *PATTERN Recognit*. 2009;42(4, SI):581-586.

16. Adler RJ, Taylor JE. *Random Fields and Geometry*. Vol 1st ed. New York: Springer-Verlag New York; 2007.

17. Sutherland DH, Kaufman KR, Wyatt MP, Chambers HG, Mubarak SJ. Double-blind study of botulinum A toxin injections into the gastrocnemius muscle in patients with cerebral palsy. *Gait Posture*. 1999;10(1):1-9.

18. Schwartz MH, Trost JP, Wervey R a. Measurement and management of errors in quantitative gait data. *Gait Posture*. 2004;20(2):196-203.

19. Sutherland DH, Davids JR. Common Gait Abnormalities of the Knee in Cerebral Palsy. *Clin Orthop Relat Res*. 1993;288:139-147.

20. Rodda JM, Graham HK, Carson L, Galea MP, Wolfe R. Sagittal gait patterns in spastic diplegia. *J Bone Joint Surg Br*. 2004;86(2):251-258.

21. Nieuwenhuys A, Papageorgiou E, Dimitriou G, Molenaers G, De Laet T, Desloovere K. Inter- and intrarater reliability of an international consensus-based gait classification system in children with cerebral palsy. *Gait Posture*. 2015;42:Supplement 1, S46.

Supporting Information

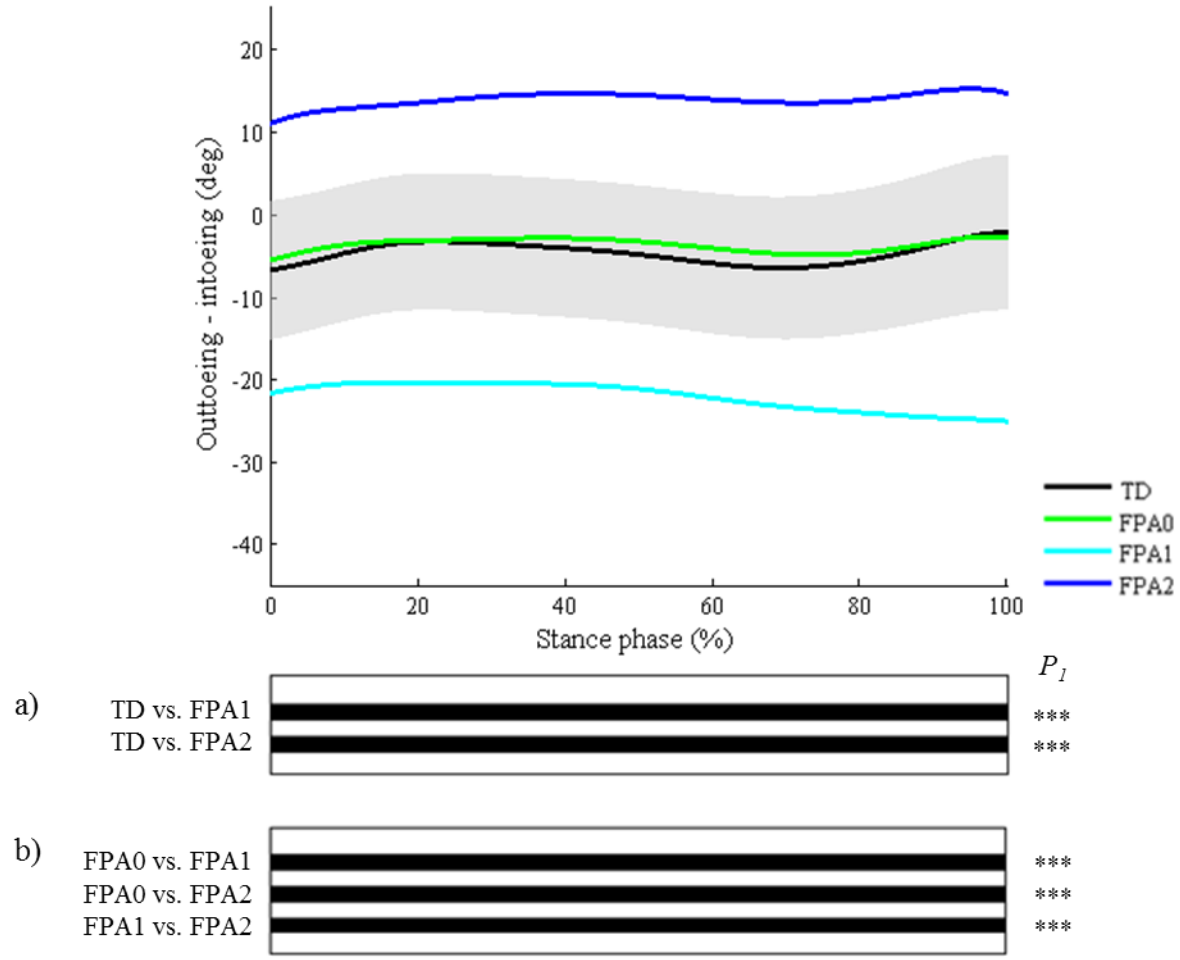


Figure S1. Top graph shows the mean kinematic angle of TD gait and of each consensus-based pattern of the foot progression angle (FPA). Black bars indicate significant gait phases during which the SPM{t} statistic exceeded the critical threshold. Panel (**a**) shows results of hypothesis 2 (i.e. unpaired t-tests, α=0.01); panel (**b**) shows results of hypothesis 3 (i.e. post-hoc unpaired t-tests, α=0.003). * $p<0.01$, ** $p<0.001$, *** $p<0.00001$. P_1 indicates the p-value of the first cluster during the stance phase.

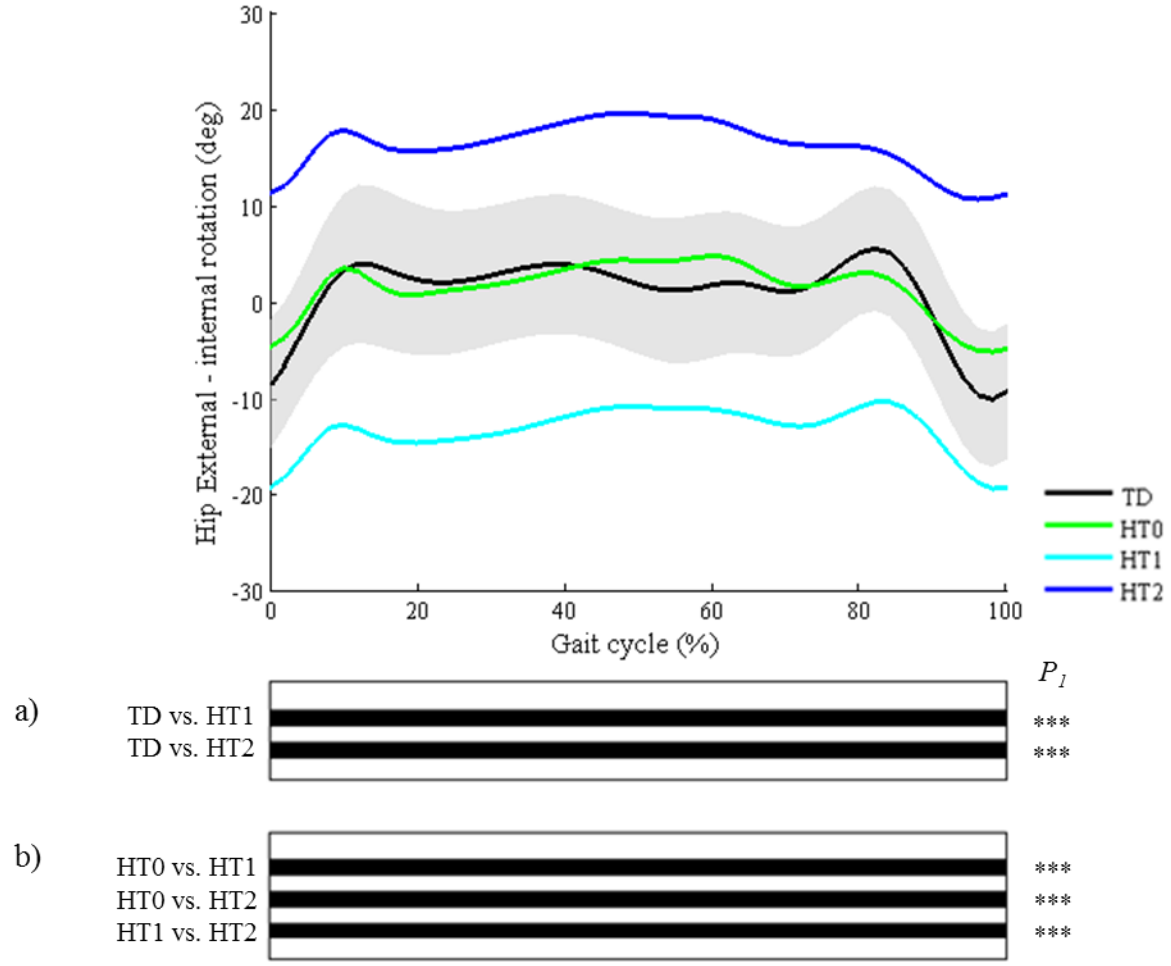


Figure S2. Top graph shows the mean kinematic angle of TD gait and of each consensus-based pattern at the level of the hip in the transverse plane (HT). Black bars indicate significant gait phases during which the SPM{t} statistic exceeded the critical threshold. Panel (**a**) shows results of hypothesis 2 (i.e. unpaired t-tests, α=0.01); panel (**b**) shows results of hypothesis 3 (i.e. post-hoc unpaired t-tests, α=0.003). * $p<0.01$, ** $p<0.001$, *** $p<0.00001$. P_1 indicates the p-value of the first cluster during the gait cycle.

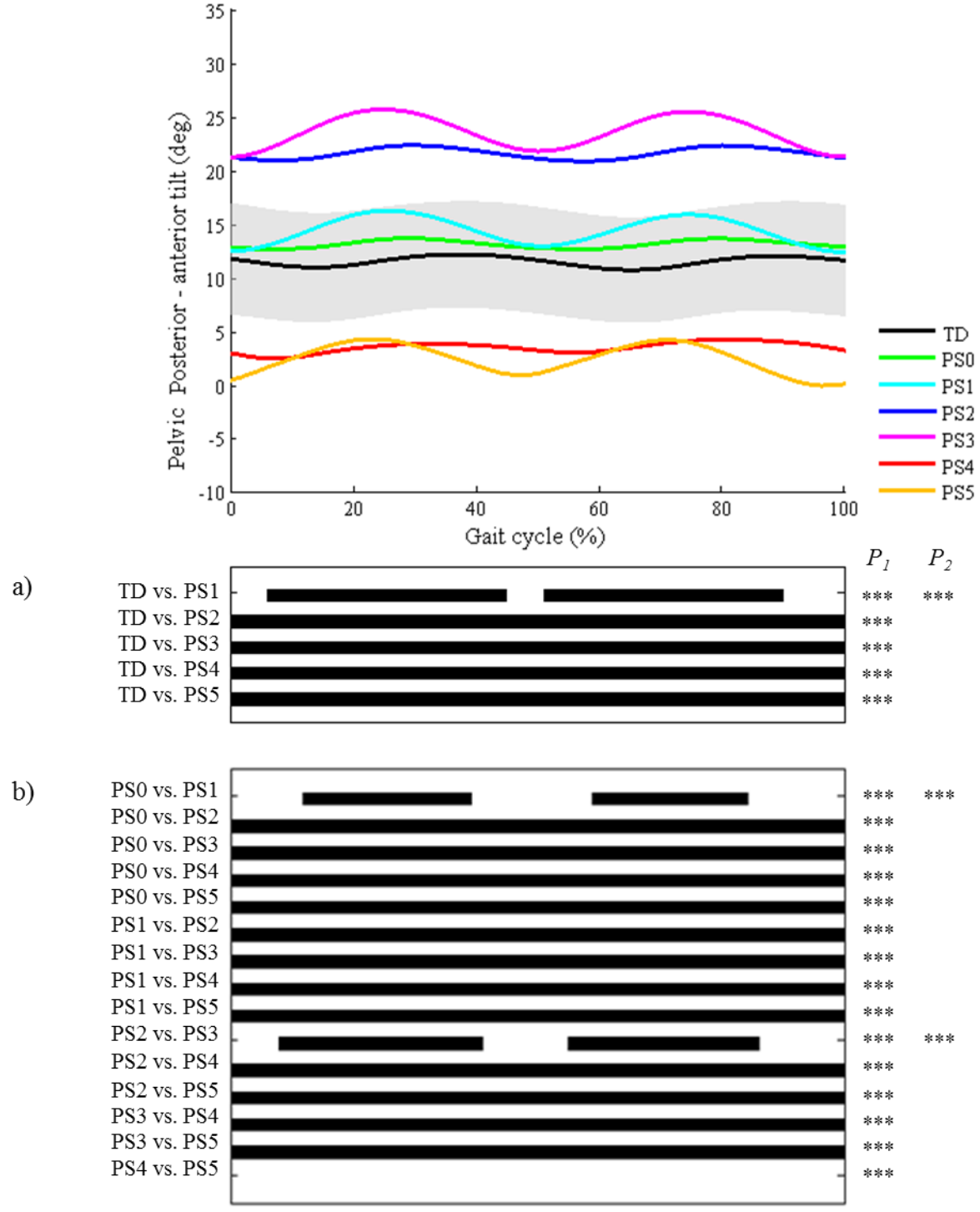


Figure S3. Top graph shows the mean kinematic angle of TD gait and of each consensus-based pattern at the level of the pelvis in the sagittal plane (PS). Black bars indicate significant gait phases during which the SPM{t} statistic exceeded the critical threshold. Panel (**a**) shows results of hypothesis 2 (i.e. unpaired t-tests, α=0.01); panel (**b**) shows results of hypothesis 3 (i.e. post-hoc unpaired t-tests, α=0.0006). * p<0.01, ** p<0.001, *** p<0.00001. P_1 indicates the p-value of the first cluster during the gait cycle, P_2 the second cluster.

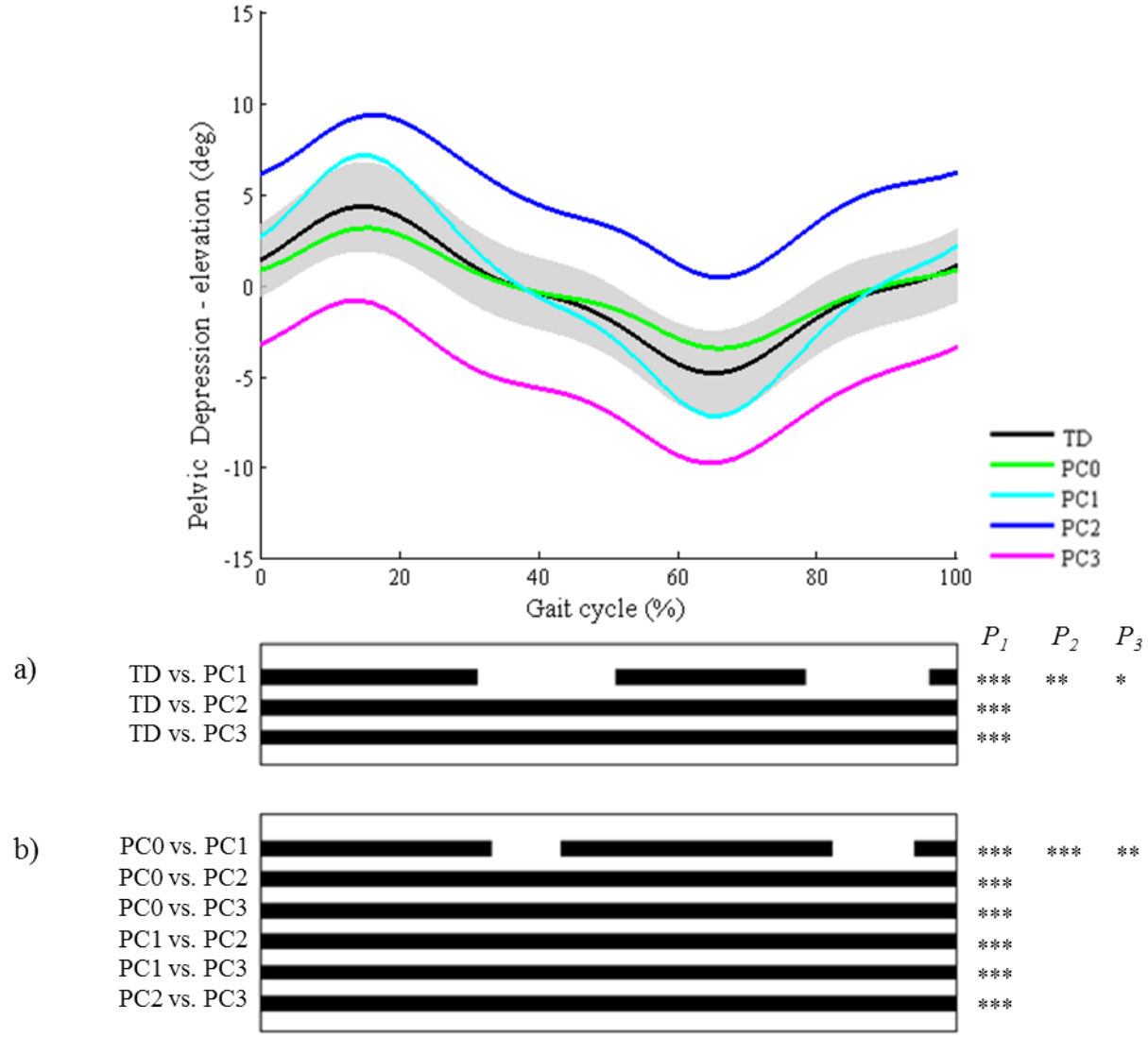


Figure S4. Top graph shows the mean kinematic angle of TD gait and of each consensus-based pattern at the level of the pelvis in the coronal plane (PC). Black bars indicate significant gait phases during which the SPM{t} statistic exceeded the critical threshold. Panel (**a**) shows results of hypothesis 2 (i.e. unpaired t-tests, α=0.01); panel (**b**) shows results of hypothesis 3 (i.e. post-hoc unpaired t-tests, α=0.002). * $p<0.01$, ** $p<0.001$, *** $p<0.00001$. P_1 indicates the p-value of the first cluster during the gait cycle, P_2 the second cluster, etc.

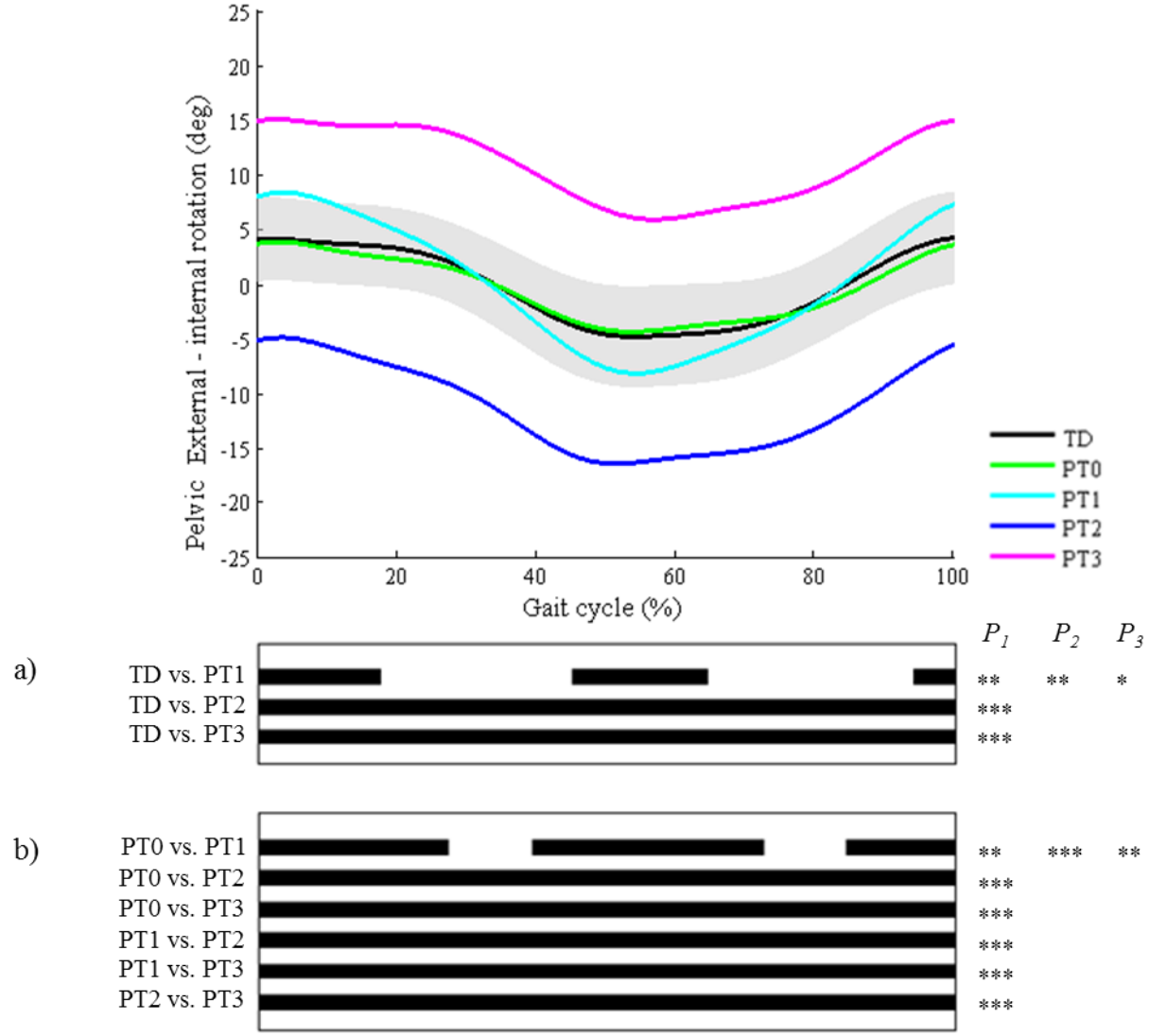


Figure S5. Top graph shows the mean kinematic angle of TD gait and of each consensus-based pattern at the level of the pelvis in the transverse plane (PT). Black bars indicate significant gait phases during which the SPM{t} statistic exceeded the critical threshold. Panel (**a**) shows results of hypothesis 2 (i.e. unpaired t-tests, α=0.01); panel (**b**) shows results of hypothesis 3 (i.e. post-hoc unpaired t-tests, α=0.002). * $p<0.01$, ** $p<0.001$, *** $p<0.00001$. P_1 indicates the p-value of the first cluster during the gait cycle, P_2 the second cluster, etc.

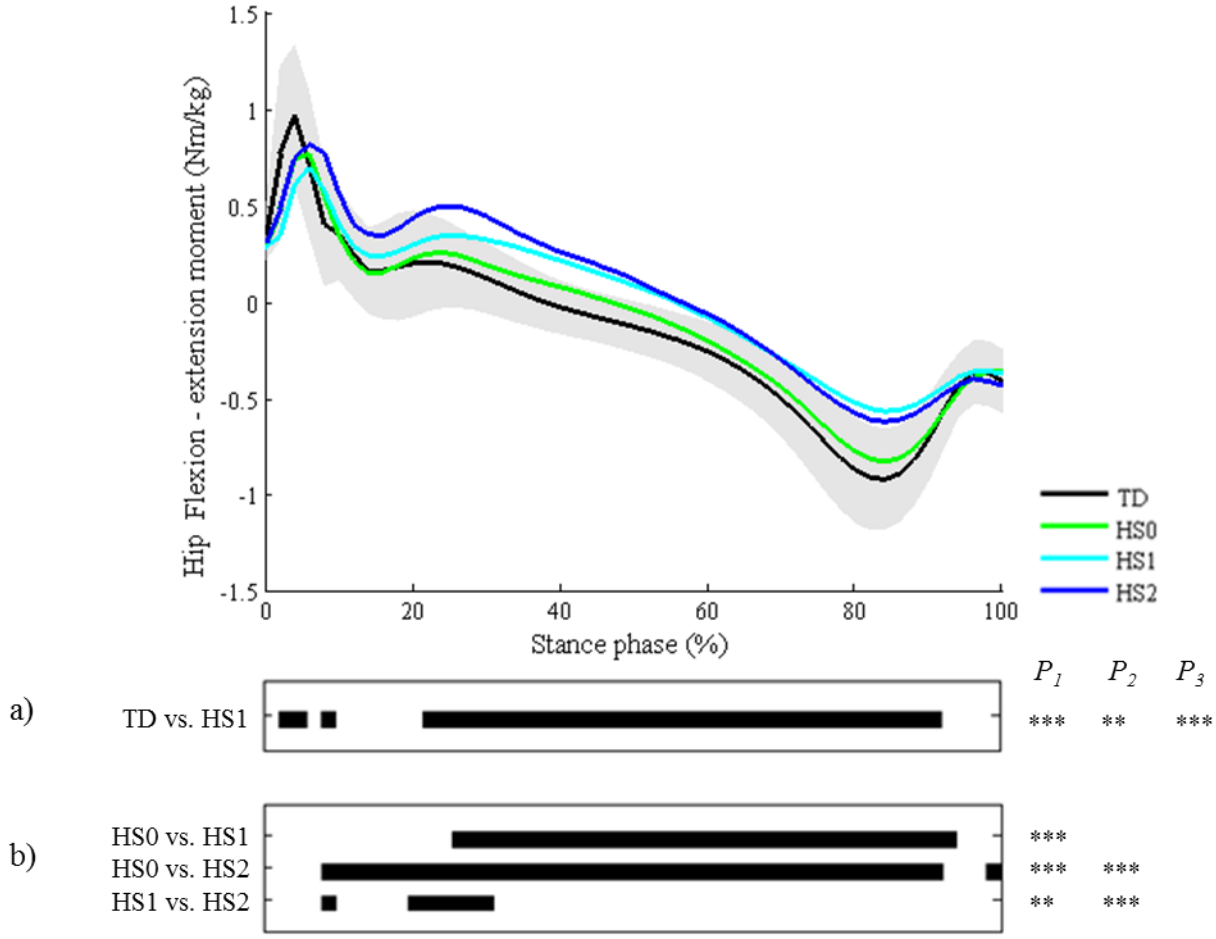


Figure S6. Top graph shows the mean kinetic angle of TD gait and of each consensus-based pattern at the level of the hip in the sagittal plane (HS). Black bars indicate significant gait phases during which the SPM{t} statistic exceeded the critical threshold. Panel (**a**) shows results of hypothesis 2 (i.e. unpaired t-tests, α=0.01); panel (**b**) shows results of hypothesis 3 (i.e. post-hoc unpaired t-tests, α=0.003). * $p<0.01$, ** $p<0.001$, *** $p<0.00001$. P_1 indicates the p-value of the first cluster during the stance phase, P_2 the second cluster, etc.

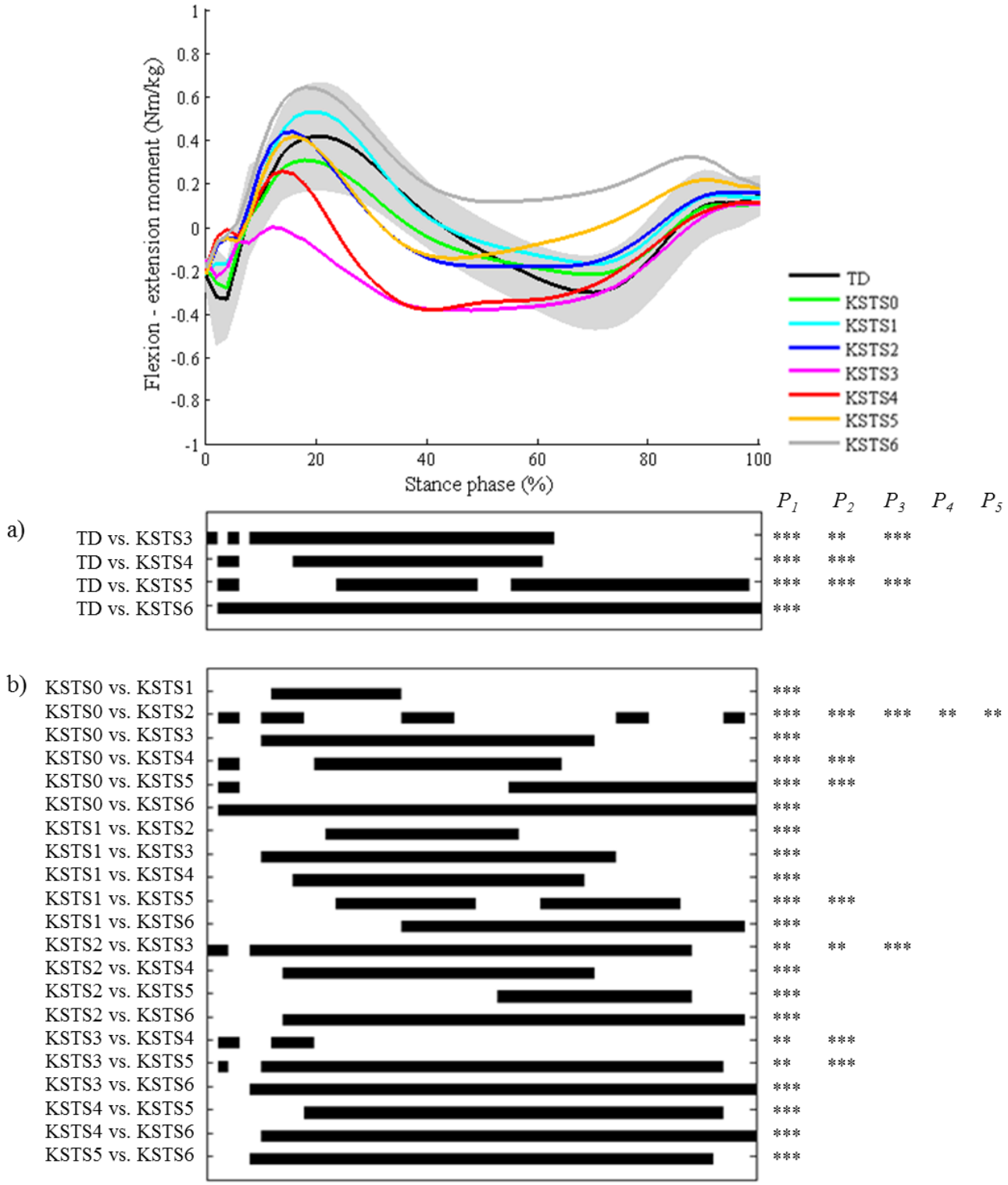


Figure S7. Top graph shows the mean kinetic angle of TD gait and of each consensus-based pattern at the level of the knee during stance phase in the sagittal plane (KSTS). Black bars indicate significant gait phases during which the SPM{t} statistic exceeded the critical threshold. Panel (**a**) shows results of hypothesis 2 (i.e. unpaired t-tests, α=0.01); panel (**b**) shows results of hypothesis 3 (i.e. post-hoc unpaired t-tests, α=0.0005). * $p<0.01$, ** $p<0.001$, *** $p<0.00001$. P_1 indicates the p-value of the first cluster during the stance phase, P_2 the second cluster, etc.

Chapter 6

Prevalence of joint patterns during gait in children with cerebral palsy is related to gross motor function, topographical classification, weakness, and spasticity

Angela Nieuwenhuys
Eirini Papageorgiou
Tinne De Laet
Guy Molenaers
Kaat Desloovere

Frontiers in Human Neuroscience [in review]

Abstract

Although several gait classifications for children with cerebral palsy (CP) have been previously defined in literature, their reliability and validity often remains poorly documented. The present study aims to provide first insight toward the construct validity and clinical relevance of 49 gait patterns in children with CP, which were recently developed during a Delphi consensus study.

A retrospective sample of convenience consisted of 286 patients with spastic CP between 3-18 years old and GMFCS level I-III. The majority of the patients were diagnosed as bilateral CP (n=166). Kinematic and kinetic trials from three-dimensional gait analysis were classified according to the definitions of the Delphi study, and one classified trial was randomly selected for each included limb (n=446). Isometric muscle weakness and spasticity were also assessed for different muscle groups acting around the hip, knee, and ankle. Subsequently, Pearson Chi square tests, Cramer's V, and adjusted standardized residuals were calculated to explore the strength and direction of the associations between the gait patterns, and the different patient-specific characteristics (i.e. age, GMFCS level, and topographical classification) and clinical symptoms (muscle weakness and spasticity).

Patient-specific characteristics showed several significant associations with the patterns of different joints, but the strength of most identified associations was weak. The results further showed that patterns with 'no or minor gait deviations' were observed most frequently in all joints except for the knee during stance phase and pelvis in the sagittal plane. These patterns with 'no or minor gait deviations' were observed significantly more often in limbs with a lower level of spasticity and good muscle strength. Several other pathological gait patterns were moderately associated with weakness and spasticity. Associations with clinical symptoms were consistently stronger for the joints in the sagittal plane, possibly because most of the evaluated muscles in this study mainly perform sagittal plane motions. Remarkably, the hip patterns in the coronal plane did not associate significantly with any of the investigated variables.

Although further validation is warranted, this study contributes to the construct validity of the gait patterns of the Delphi study, by demonstrating their ability to distinguish between clinically relevant subgroups in CP.

Introduction

Cerebral palsy (CP) is the result of a pre- or post-natal lesion in the developing brain of a fetus or child, primarily affecting motor behavior. The heterogenic clinical presentation of CP is emphasized, not only because of the numerous potential differences in timing, location, severity, and nature of brain lesions, but also because it is continuously altered by a maturing brain, musculoskeletal growth, and treatment[1]. For epidemiological, treatment-related, and many other reasons, it is therefore important to identify relevant subgroups within the CP population. Several important categorizations of subgroups in CP have been reported before. For instance, the Gross Motor Function Classification Scale (GMFCS) and the Manual Ability Classification Scale are used to classify the severity of lower and upper limb motor function impairment[2,3]. Because of the complex interaction between primary and secondary motor symptoms in CP, for example between spasticity and muscle contractures, gait pathology varies a lot between patients. Hence, functional classifications such as the GMFCS often cannot cover all deviations[4].

In literature, several gait classifications have been defined based on three-dimensional gait analysis data (i.e. kinematics, kinetics or muscle activation data)[5–8]. Their utility in clinical practice is hindered because the psychometric properties of reliability and validity are often not established[5]. Recently, a new overview of gait patterns for all ambulatory children with spastic CP has been described, covering the wide range of gait deviations in the relevant lower limb joints across the three anatomical planes[9]. Via a Delphi consensus study, an expert panel defined 49 gait patterns for the pelvis, hip, knee, and ankle joints in the sagittal, coronal, and transverse plane. Previous research showed that the created classification can be reliably used, even by inexperienced clinicians[10]. However, their construct validity and relevance for clinical and research practice has not yet been examined.

The construct validity can be assessed, for instance by comparing the gait classification with a criterion classification[11], or by assessing its relationships with scores of other instruments. Previous research has already shown the relevance of establishing the relation between specific gait features and other variables such as topographical classification, age, preceding treatments, and clinical measurements[12,13]. Further, Rozumalski et al.[14] investigated how different crouch gait patterns, which were determined via k-means cluster analysis, were characterized by range of motion, muscle strength, and spasticity. Dobson et al.[15] reported on the construct validity of the Winters classification, by showing how the distribution of the

patterns was associated with other validated classifications such as the Gross Motor Function Classification Scale[2] (GMFCS) and Functional Mobility Scale[16]. By providing evidence that the classification can make a distinction between relevant subgroups in CP, its usefulness and validity can be demonstrated.

The present study aims to provide first insight toward the construct validity and clinical relevance of the aforementioned consensus-based gait patterns in CP[9]. The prevalence of the patterns and their association with other patient-specific characteristics and clinical symptoms is explored in an extended cohort of children with CP. It is hypothesized that the prevalence of the patterns is associated with age, topographical classification, GMFCS level, and previous treatment. The study also examines how specific gait patterns are characterized by weakness and spasticity. It is hypothesized that pelvis and hip patterns are associated in particular with the severity of weakness or spasticity in muscle groups that have a function around the pelvis and hip joint. Analogous to the previous hypothesis, knee and ankle patterns are expected to associate with the presence of weakness or spasticity in the muscles acting at the knee and ankle respectively.

Methodology

Patient recruitment and data collection

This study was approved by the Medical Ethical Committee of University Hospitals Leuven (s56036). An extended retrospective convenience sample was available from the database of the hospital, comprising gait analysis sessions that were obtained for research or clinical purposes between November 2001 and August 2015. The sample contained a total of 459 sessions (from 356 children), which were all screened for the following inclusion criteria: (a) a diagnosis of unilateral or bilateral CP (b) predominantly spastic type of CP (c) 3-18 years of age, (d) GMFCS-level I-III, and (e) the availability of at least two good quality kinematic gait trials from three-dimensional gait analysis.

Instrumented gait analysis

Standardized three-dimensional gait analyses were performed using ten to fifteen VICON motion camera's (Vicon Motion Systems, Oxford, UK) and two AMTI force plates (Advanced Mechanical Technology Inc., Watertown, MA, USA). Reflecting markers were placed on anatomical landmarks of the patient according to the Plug-In-Gait marker model

and patients were instructed to walk barefoot and at a self-selected speed on a 10m-walkway. Nexus software was used to define gait cycles and to estimate joint angles and joint moments in the three anatomical planes. For each kinematic trial, one gait cycle per side (left and right) was identified. Both the left and right side were included for all patients with bilateral CP. For patients with unilateral CP, only the affected body side was selected for analysis. All available kinematic and kinetic trials were visually screened and trials with artifacts, signs of inaccurate marker placement, or trials that were not representative of a patient's gait (outliers), were excluded so that only trials of good quality remained. The remaining trials, 1719 in total, were classified by a clinician who was experienced with the gait patterns (AN or EP). As a result, for each gait analysis session, one to seven trials per side per patient were classified. Subsequently, for each included session, one classified trial was randomly selected per side, unless a pattern with a very low prevalence in the database was present (<10% of 1719 trials), in which case that trial was given priority. The interrater reliability between both raters was previously shown to be almost perfect (overall percentage of agreement=90%, kappa=0.86, confidence interval=078-0.94). Table 1 shows the different joints that were classified as well as a concise description of the patterns per joint.

One gait analysis session was selected for each patient. Sessions were excluded if a patient had undergone Botulinum toxin type A treatment less than 180 days or surgery (i.e. single event multilevel surgery or selective dorsal rhizotomy) less than 365 days before the date of the gait analysis session. In case more than one session was still available for a patient, preference was given to the earliest pre-treatment session with the least amount of missing data from the clinical examination.

Clinical examination of weakness and spasticity

Gait analysis sessions were preceded by a clinical examination during which muscle strength and muscle tone were evaluated. Isometric muscle strength was assessed by experienced physiotherapists using the manual muscle testing scale (MMT)[17,18]. The MMT is scored on a six-point ordinal scale (scores range from 0-5) and it differentiates between a palpable contraction and a motion against gravity or against resistance. The maximum score of 5 indicates that a patient can move for the full range of motion against gravity and maximum resistance, whereas a score of 0 indicates that no contraction can be palpated. Isometric strength was assessed and scored for the following muscle groups: hip flexors, extensors, adductors, and abductors; knee flexors and extensors; ankle dorsi- and plantar flexors, and the

muscle groups performing ankle inversion and eversion. In addition, muscle spasticity was evaluated using the Modified Ashworth Scale (MAS)[19], which is also a six-point ordinal scale (scores: 0, 1, 1+, 2, 3, 4), that measures the extent of increase in muscle tone in combination with the feeling of a catch during the stretch of a passive muscle group through the full range of motion. The maximum score of 4 indicates that the evaluated muscle or muscle group is rigid and no motion is possible, whereas a score of 0 indicates a normal muscle tone. MAS scores were collected for the hip flexors, short adductors, and long adductors; for the hamstrings and rectus femoris muscles at the level of the knee; and for the gastrocnemius, soleus, and tibialis posterior muscles at the level of the ankle joint.

Because of the high number of muscles that were evaluated during the clinical examination and because of the explorative nature of the study, it was decided to group the muscles according to the joints around which they have their main function, such that the hip, knee, and ankle joint were characterized by one score for muscle weakness and one score for spasticity. For instance, the highest MAS score between the gastrocnemius, soleus, and tibialis posterior muscles was selected to represent the severity of spasticity around the ankle joint. The involved multidisciplinary team advized to select the most severe score for weakness (i.e. lowest score) and spasticity (i.e. highest score) at the level of each joint because of two reasons: on the one hand, the muscles most affected by weakness or spasticity were considered to have a larger influence on pathological gait deviations. On the other hand, the selection of the most severe score per joint, instead of averaged values or summation of muscle-specific scores, ensured that the impact of weakness or spasticity would not be filtered out (which might be expected if the average of the joint sub-scores was used). In addition, the clinical examination data was characterized by missing data as a result of the retrospective nature of the study. By selecting the most severe score per joint, the sample size of the study would not be reduced, which was expected to happen if the muscle-specific scores were summed. The influence of these missing data on the results was expected to be negligible, as the median percentage of missing data per MAS or MMT variable was 0.44% (range 0%-5.6%).

Data selection and statistical analysis

Descriptive statistics and cross-tables were used to describe the frequency distributions for all gait patterns, as well as for the following patient-specific characteristics and clinical symptoms: age, GMFCS level, previous orthopedic surgery, topographical classification

(unilateral vs. bilateral CP), and clinical examination scores (i.e. weakness of the muscles around the hip, knee, and ankle; spasticity of the muscles around the hip, knee, and ankle). Age was further categorized into three groups using the 25^{th} and 75^{th} percentile as cut-off values. These categories will further be referred to as the 'youngest patients' (patients until 7.5 years old), 'medium aged patients' (patients from 7.5-12.5 years old), and 'oldest patients' (patients over 12.5 years old).

Pearson Chi-square tests (χ^2) were performed to investigate if the distribution of the patient-specific characteristics and clinical symptoms were significantly related to the distribution of the gait patterns at the level of each joint (α=0.05). To allow for a valid interpretation of χ^2, a sufficiently large sample size is required and expected frequencies below n=5 can only be accepted in less than 20% of the cells of the cross-tables[20]. If this condition was not met, categories of a variable were combined, but only if merging those categories was clinically meaningful (e.g. Scores 4 and 5 of the MMT were often combined, both scores indicating that the patient could move against moderate to heavy resistance). If significant associations were identified, the strength of the association was evaluated using Cramer's V, which was interpreted to be weak, moderate or strong, depending on the degrees of freedom (Table S1)[21]. Subsequently, adjusted standardized residuals (ASR) were examined to explore the direction of significant associations. ASRs can identify significant combinations of specific categories of two variables that contributed stronger to the identified association than other combinations of categories. Because ASRs follow a normal distribution with mean '0' and standard deviation '1', ASR values larger than -2 or +2 indicate that the frequency count in a particular cell is respectively significantly smaller or higher than would be expected if the two variables were unrelated ($p<0.05$).

Table 1. Brief definition of all joint patterns during gait and their prevalence in the selected limbs (N=446) from the patient population.

SAGITTAL PLANE	N (%)
Pelvis	
PS0 - Normal pelvic motion/posture – no or minor gait deviations	88 (19.7)
PS1 - Increased range of motion	130 (29.1)
PS2 - Increased anterior tilt on average	67 (15.0)
PS3 - Increased anterior tilt and increased range of motion	157 (35.2)
PS4 - Decreased anterior tilt (posterior tilt)	1 (0.2)
PS5 - Decreased anterior tilt (posterior tilt) and increased range of motion	3 (0.7)
Hip	
HS0 - Normal hip motion – no or minor gait deviations	229 (51.3)
HS1 - Hip extension deficit	136 (30.5)
HS2 - Continuous excessive hip flexion	81 (18.2)
Knee during stance	
KSTS0 - Normal knee motion during stance – no or minor gait deviations	56 (12.6)
KSTS1 - Increased knee flexion at initial contact	33 (7.4)
KSTS2 - Increased knee flexion at initial contact and earlier knee extension movement	89 (20.0)
KSTS3 - Knee hyperextension	38 (8.5)
KSTS4 - Knee hyperextension and increased knee flexion at initial contact	53 (11.9)
KSTS5 - Increased flexion in midstance and internal flexion moment present	100 (22.4)
KSTS6 - Increased flexion in midstance and internal extension moment present	77 (17.3)
Knee during swing	
KSWS0 - Normal knee motion during swing – no or minor gait deviations	140 (31.4)
KSWS1 - Delayed peak knee flexion	103 (23.1)
KSWS2 - Increased peak knee flexion	50 (11.2)
KSWS3 - Increased and delayed peak knee flexion	42 (9.4)
KSWS4 - Decreased peak knee flexion	53 (11.9)
KSWS5 - Decreased and delayed peak knee flexion	58 (13.0)
Ankle during stance	
ASTS0 - Normal ankle motion during stance – no or minor gait deviations	164 (36.8)
ASTS1 - Horizontal second ankle rocker	133 (29.8)
ASTS2 - Reversed second ankle rocker	53 (11.9)
ASTS3 - Equinus gait	22 (4.9)
ASTS4 - Calcaneus gait	74 (16.6)
Ankle during swing	
ASWS0 - Normal ankle motion during swing – no or minor gait deviations	165 (37.0)
ASWS1 - Insufficient prepositioning in terminal swing	39 (8.7)
ASWS2 - Continuous plantarflexion during swing (drop foot)	94 (21.1)
ASWS3 - Excessive dorsiflexion during swing	148 (33.2)

Table 1. Continued.

CORONAL PLANE	N (%)
Pelvis	
PC0 - Normal pelvic motion/posture – no or minor gait deviations	225 (50.4)
PC1 - Increased pelvic range of motion	135 (30.3)
PC2 - Continuous pelvic elevation	34 (7.6)
PC3 - Continuous pelvic depression	52 (11.7)
Hip	
HC0 - Normal hip motion – no or minor gait deviations	278 (62.3)
HC1 - Excessive hip abduction in swing	87 (19.5)
HC2 - Continuous excessive hip abduction	52 (11.7)
HC3 - Continuous excessive hip adduction	29 (6.5)
TRANSVERSE PLANE	
Pelvis	
PT0 - Normal pelvic motion/posture – no or minor gait deviations	204 (45.7)
PT1 - Increased pelvic range of motion	136 (30.5)
PT2 - Excessive pelvic external rotation during the gait cycle	66 (14.8)
PT3 - Excessive pelvic internal rotation during the gait cycle	40 (9.0)
Hip	
HT0 - Normal hip motion – no or minor gait deviations	338 (75.8)
HT1 - Excessive hip external rotation during the gait cycle	34 (7.6)
HT2 - Excessive hip internal rotation during the gait cycle	74 (16.6)
Foot progression angle	
FPA0 - Normal foot progression angle – no or minor gait deviations	279 (62.6)
FPA1 - Outtoeing	73 (16.4)
FPA2 - Intoeing	94 (21.1)

Described deviations such as increased or excessive joint angles refer to deviations which are more than one standard deviation away from a reference database of typically developing children. A more detailed description of the patterns is available in Nieuwenhuys et al.[9].

Results

Description of experimental patient population

After the data selection process, the experimental sample consisted of 286 patients with spastic CP of which the majority had a diagnosis of bilateral CP (n=166) and the median age was 10.2 years (Table 2). Gait analysis sessions of patients who had undergone previous orthopedic surgery were collected after a median of approximately 2 years (interquartile range: 1 year and 3 months – 5 years and 6 months). Because both sides could be included for

the majority of the patients with bilateral CP, a total of 446 limbs were used for the statistical analyses of side-specific variables (i.e. 'previous surgery', spasticity and weakness scores).

Table 3 presents the frequency distribution of the spasticity and weakness scores around the hip, knee, and ankle joint. The muscles acting around the hip were least affected by spasticity, with 48.5% of all limbs classified as MAS 0 or 1. On the contrary, muscles around the ankle joint were most severely affected by spasticity, with 42.7% of all limbs classified as MAS 2, 3, or 4. The weakest muscle groups were also those with their main function around the ankle, with 16.3% of all limbs classified as MMT 0 or 1 as opposed to 1.6% and 0% for the same MMT scores at the hip and knee joint.

Table 1 presents the prevalence of the 49 patterns. For all joints except for the knee during stance and pelvis in the sagittal plane, the pattern with 'no or minor gait deviations' was most prevalent, indicating that patients mostly remained within one standard deviation from the mean of an age-matched group of typically developing children. Pathological patterns that were observed most frequently in the proximal joints were 'increased pelvic anterior tilt and increased range of motion' (35.2%), 'hip extension deficit' (30.5%), and 'increased pelvic range of motion' in the sagittal (29.1%), coronal (30.3%) and transverse (30.5%) plane. For the distal joints, the patterns 'excessive ankle dorsiflexion during swing' (33.2%), 'horizontal second ankle rocker during stance' (29.8%), 'delayed peak knee flexion during swing' (23.1%), and 'excessive knee flexion and internal flexion moment during stance' (22.4%) were most frequently observed. Because the prevalence of 'decreased pelvic anterior tilt' (0.2%) and 'decreased pelvic anterior tilt and increased range of motion' (0.7%) was extremely low, both patterns needed to be excluded from further statistical analyses.

Tables 4 and 5 report the results of all χ^2 analyses, which established the associations between the distribution of the gait patterns and the patient-specific variables, previous surgery, spasticity, and weakness. Because many significant associations were identified, only the directions of significant moderate associations, where the ASR reached a value larger than 2, are discussed in detail (Figures 1-6). Detailed information on the direction of significant weak associations (ASRs) is available in tables S2-S6 in the supplementary material.

Table 2. Patient characteristics (N=286).

	N (%)	
Gender		
Male	165 (57.7)	
Female	121 (42.3)	
Diagnosis		
Bilateral CP	166 (58.0)	
Unilateral CP	120 (42.0)	
GMFCS		
Level I	172 (60.1)	
Level II	89 (31.1)	
Level III	25 (8.7)	
Previous orthopedic surgery		
Yes	55 (19.2)	(n=100 limbs)
No	231 (80.8)	(n=346 limbs)
Number of previous Botulinum Toxin type A		
None	111 (38.8)	(n=159 limbs)
One or two	104 (36.4)	(n=155 limbs)
Three or more	71 (24.8)	(n=132 limbs)
Weight (mean (SD), in kg)	34.3 (14.8)	
Height (mean (SD), in cm)	137.6 (19.7)	
Age at time of gait analysis (median (IQR), in years)	10.2 (7.5-12.5)	

SD = standard deviation; IQR = interquartile range.

Table 3. Prevalence and distribution of MAS and MMT scores for the muscles around the hip, knee, and ankle joint in the selected limbs (N=446) from the patient population.

	MAS score (N (%))					
	0	**1**	**1+**	**2**	**3**	**4**
Hip	93 (20.9)	123 (27.6)	130 (29.1)	98 (22.0)	2 (0.4)	0 (0.0)
Knee	22 (4.9)	118 (26.5)	153 (34.3)	142 (31.8)	11 (2.5)	0 (0.0)
Ankle	9 (2.0)	46 (10.3)	196 (43.9)	164 (36.8)	26 (5.8)	5 (1.1)
	MMT score (N (%))					
	0	**1**	**2**	**3**	**4**	**5**
Hip	0 (0.0)	7 (1.6)	33 (7.4)	231 (51.8)	162 (36.3)	13 (2.9)
Knee	0 (0.0)	0 (0.0)	14 (3.1)	191 (42.8)	221 (49.6)	20 (4.5)
Ankle	5 (1.1)	68 (15.2)	85 (19.1)	189 (42.4)	83 (18.6)	16 (3.6)

If less than 50 limbs were classified in a particular category of the MMT or MAS scale, the expected frequencies in the cross-tables were generally too low to allow a valid interpretation of χ^2, especially for analyses in combination with joints that have a high number of patterns (e.g. knee during stance (n=7)). Therefore, darker shaded categories were merged at the level of each joint, all indicating a lower level of spasticity or a higher level of muscle weakness. Lightly shaded areas were merged at the level of each joint, indicating a higher level of spasticity and a lower level of muscle weakness.

Table 4. Pearson chi squared analyses (χ^2) and Cramer's V (*V*) identified significantly weak, moderate, and strong associations between the sagittal plane joint patterns and patient-specific characteristics, previous surgery, spasticity, and weakness.

	PS[b]			HS			KSTS			KSWS			ASTS			ASWS		
	χ^2		*V*	χ^2		*V*	χ^2		*V*	χ^2		*V*	χ^2		*V*	χ^2		*V*
N = 286 patients																		
Uni-/bilateral CP	7.77		0.17	8.84	*	0.18	24.69	**	0.29	27.46	***	0.31	5.83		0.14	20.66	**	0.27
Age	13.21	*	0.15	11.03	*	0.14	16.95		0.17	37.08	***	0.26	28.02	**	0.22	9.02		0.13
GMFCS	38.96	***	0.26	30.49	***	0.23	64.70[a]	***	0.34	53.73[a]	***	0.31	27.00[a]	*	0.22	10.31		0.13
N = 446 limbs																		
Previous surgery	8.26	*	0.14	8.83	*	0.14	14.40	*	0.18	1.05		0.05	18.70	*	0.21	55.71	***	0.35
MAS Hip joint	68.51	***	0.23	41.95	***	0.22	81.37	***	0.25	149.48	***	0.33	44.60	***	0.18	14.16		0.10
MAS Knee joint	44.23	***	0.22	27.41	***	0.18	71.86	***	0.28	91.68	***	0.32	29.64	**	0.18	18.47	*	0.14
MAS Ankle joint	29.12	***	0.26	4.07		0.10	39.30	***	0.30	67.69	***	0.39	42.28	***	0.31	17.20	*	0.20
MMT Hip joint	52.18	***	0.34	30.25	***	0.26	48.80	***	0.33	51.31	***	0.34	9.35		0.15	12.82	*	0.17
MMT Knee joint	57.67	***	0.36	35.44	***	0.28	36.51	***	0.29	72.23	***	0.40	18.91	*	0.21	10.33	*	0.15
MMT Ankle joint	79.96	***	0.25	38.31	***	0.21	59.66	***	0.21	78.05	***	0.24	28.85	*	0.15	33.43	**	0.16

* $p<0.05$; ** $p<0.001$; *** $p<0.0001$; χ^2 = Pearson chi squared; *V*=Cramer's V, indicating significantly weak (light grey), moderate (darker grey), and strong (dark grey) associations based on degrees of freedom (Table S1); [a] results should be interpreted with caution because >20% of cells had expected frequencies lower than n=5; [b] N=282 patients and N=442 limbs due to exclusion of PS4 and PS5.

Table 5. Pearson chi squared analyses (χ^2) and Cramer's V (*V*) identified significantly weak and moderate associations between the coronal and transverse plane joint patterns and patient-specific characteristics, previous surgery, spasticity, and weakness.

	PC			HC			PT			HT			FT		
	χ^2		*V*	χ^2		*V*	χ^2		*V*	χ^2		*V*	χ^2		*V*
N = 286															
Uni-/bilateral CP	24.92	***	0.30	2.42		0.09	26.49	***	0.30	3.10		0.10	14.56	*	0.23
Age	13.63	*	0.15	4.89		0.09	4.43		0.09	2.88		0.07	11.46	*	0.14
GMFCS	10.02		0.13	17.28[a]	*	0.17	19.42	*	0.18	12.71[a]	*	0.15	7.60		0.12
N = 446															
Previous surgery	8.38	*	0.14	2.29		0.07	2.71		0.08	10.25	*	0.15	2.03		0.07
MAS Hip joint	23.84	*	0.13	3.18		0.05	15.98		0.11	28.79	***	0.18	21.75	*	0.16
MAS Knee joint	19.51	*	0.15	1.84		0.05	16.97	*	0.14	15.31	*	0.13	10.70	*	0.11
MAS Ankle joint	6.32		0.12	5.24		0.11	5.07		0.11	8.94	*	0.14	4.40		0.10
MMT Hip joint	12.64	*	0.17	1.44		0.06	11.39	*	0.16	9.31	*	0.14	5.53		0.11
MMT Knee joint	9.26	*	0.14	3.82		0.09	5.74		0.11	16.61	**	0.19	7.42	*	0.13
MMT Ankle joint	14.53		0.10	10.12		0.09	28.51	*	0.15	23.61	*	0.16	13.49	*	0.12

* $p<0.05$; ** $p<0.001$; *** $p<0.0001$; χ^2 = Pearson chi squared; *V*=Cramer's V, indicating weak (light grey) and moderate (darker grey) associations based on degrees of freedom (Table S1); [a] results should be interpreted with caution because >20% of cells had expected frequencies lower than n=5.

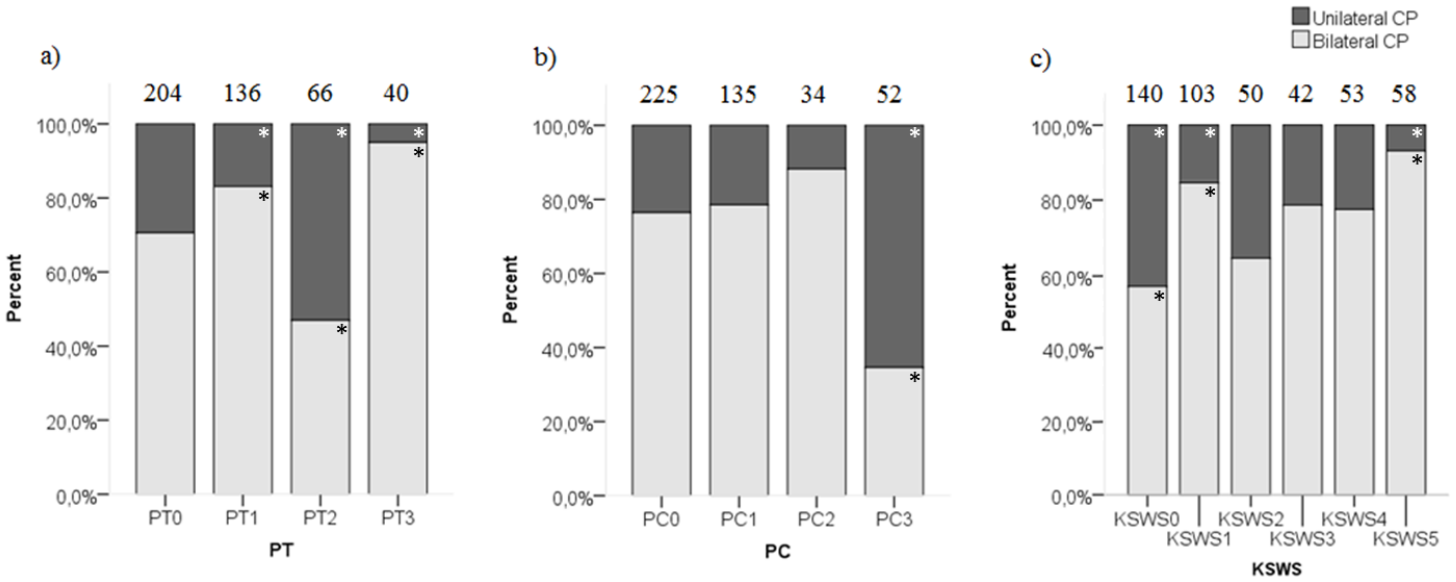


Figure 1. Topographical classification associated moderately with **a)** pelvis patterns in transverse plane (PT) **b)** pelvis patterns in coronal plane (PC) and **c)** knee patterns during swing (KSWS). * indicates that a pattern was observed significantly more or less frequently in children with unilateral or bilateral CP ($p<0.05$). Specific ASRs are available in Table S3, S5-6. Numbers on top of each bar represent the number of patients that were classified into that pattern.

Relations with patient-specific characteristics (N=286)

Topographical classification related moderately with the pelvic patterns in the transverse plane (Cramer's V=0.30, p<0.0001) and coronal plane (Cramer's V=0.30, p<0.0001) as well as with the knee patterns during swing (Cramer's V=0.31, p<0.0001) in the sagittal plane (Figure 1). Patients with unilateral CP were observed more often than expected with 'excessive pelvic external rotation', 'pelvic depression', and 'minor gait deviations' in the knee during swing phase. In addition, patients with bilateral CP were classified more often with 'increased pelvic range of motion' in the transverse plane, and 'delayed peak knee flexion' during swing.

Age showed moderate associations with the knee patterns during swing (Cramer's V=0.26, p<0.0001) and ankle patterns during stance (Cramer's V=0.22, p<0.001) in the sagittal plane (Figure 2). A 'horizontal' or 'reversed second ankle rocker' was observed significantly more often in the youngest patients, whereas the oldest patients were more often classified as 'calcaneus gait' or with 'minor gait deviations'. The youngest patients also showed more often a 'delayed peak knee flexion' or a 'delayed and increased peak knee flexion' during swing.

GMFCS level was moderately associated with the patterns of the pelvis (Cramer's V=0.26, p<0.0001) and hip (Cramer's V=0.23, p<0.0001) in the sagittal plane (Figure 3). Moderate associations were also found for the knee during stance and swing, as well as the ankle during stance. However, the results of these χ^2 analyses should be interpreted with caution due to the low number of patients classified as GMFCS level III in combination with pathological patterns that showed a low prevalence (e.g. equinus gait (4.9%)). In general, patients with GMFCS level I were observed significantly more often in the patterns with 'minor gait deviations' for the pelvis, hip, knee, and ankle joints in the sagittal plane. Patients with GMFCS levels II and III also displayed the patterns 'hip extension deficit' and 'increased pelvic anterior tilt' significantly more often than expected.

Relations with side-specific variables and clinical symptoms (N=446)

Previous surgery was moderately associated with the ankle patterns during swing (Cramer's V=0.35, p<0.0001; Figure 4). The categories that mainly contributed to this association were the higher frequency of 'excessive dorsiflexion during swing' in combination with limbs that had undergone previous surgery.

The hip in the coronal plane was the only joint not associated with weakness or spasticity (Table 5). Further, only weak associations were identified for all joints in the coronal and transverse plane. Even though the associations were all weak, it was notable that the pattern ‘excessive hip internal rotation’ was observed significantly more often in combination with higher levels of spasticity (MAS 2, 3, or 4) and weakness (MMT 0, 1, 2, or 3) for the muscles acting around the hip, knee, and ankle (Table S6).

In the sagittal plane, spasticity scores for muscles around the hip were moderately associated with the pelvis and hip patterns in the sagittal plane (Cramer’s V=0.23 and 0.22 respectively, both $p<0.0001$). Weakness at the level of the hip was moderately associated with the sagittal pelvis patterns (Cramer’s V=0.34, $p<0.0001$), and weakly associated with the sagittal hip patterns (Cramer’s V=0.26, $p<0.0001$; Figure 5). The pattern with ‘minor gait deviations’ in both the pelvis and hip joint was observed significantly more often in limbs with few signs of spasticity (MAS scores 0, 1) or weakness (MMT scores 4, 5). On the other hand, pathological patterns such as ‘increased pelvic anterior tilt and increased range of motion’ or ‘continuous excessive hip flexion’ were mainly observed in limbs that were markedly affected by spasticity (MAS 1+, 2, 3, 4) or weakness (MMT 0, 1, 2, 3).

Severity of spasticity around the knee joint was moderately associated with the knee patterns both during stance and swing (Cramer’s V=0.28 and 0.32 respectively, both $p<0.0001$; Figure 6). A moderate association was also identified between weakness scores at the level of the knee and the knee patterns during swing (Cramer’s V=0.40, $p<0.0001$). For the knee patterns during swing, it was apparent that all patterns with the feature ‘delayed peak knee flexion’ (KSWS1, KSWS3, KSWS5; Figure 6) were observed significantly more often in combination with higher levels of spasticity (MAS 2, 3, 4) and weakness (MMT 0, 1, 2, 3). For the knee patterns during stance, ‘minor gait deviations’ and ‘increased knee flexion at initial contact’ were mainly observed in limbs with few signs of spasticity (MAS 0, 1) or weakness (MMT 4,5). Limbs with higher levels of spasticity (MAS 2, 3, 4) or weakness (MMT 0, 1, 2, 3) were classified more often than expected as ‘increased knee flexion at initial contact and knee hyperextension’ as well as ‘increased flexion during midstance and internal flexion moment present’.

Spasticity at the level of the ankle was moderately associated with the ankle patterns during stance (Cramer’s V=0.31, $p<0.0001$; Figure 4), and weakly associated with the ankle patterns during swing (Cramer’s V=0.20, $p=0.001$). The patterns ‘equinus gait’ and ‘reversed second

ankle rocker' were mainly observed in combination with marked signs of spasticity (MAS 2, 3, 4). Weakness at the level of the ankle was weakly associated with the ankle patterns both during stance and swing (Cramer's V=0.15 and 0.16 respectively, both p<0.01).

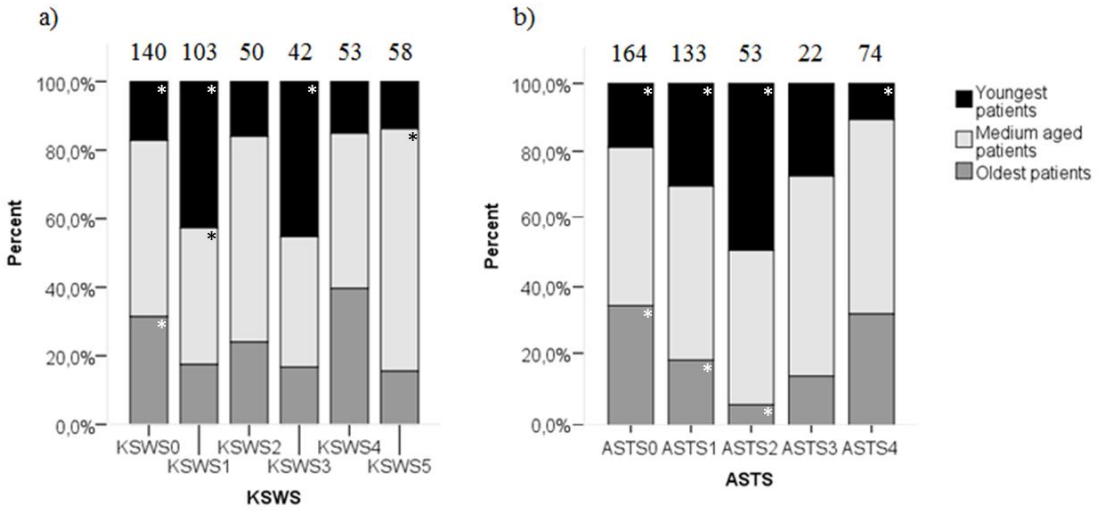


Figure 2. Age associated moderately with the distribution of **a)** knee patterns during swing (KSWS) and **b)** ankle patterns during stance (ASTS). * indicates that a pattern was present significantly more or less frequently in the youngest, medium aged, or oldest patients (p<0.05). Specific ASRs are available in Table S3-4. Numbers on top of each bar represent the number of patients that were classified into that pattern.

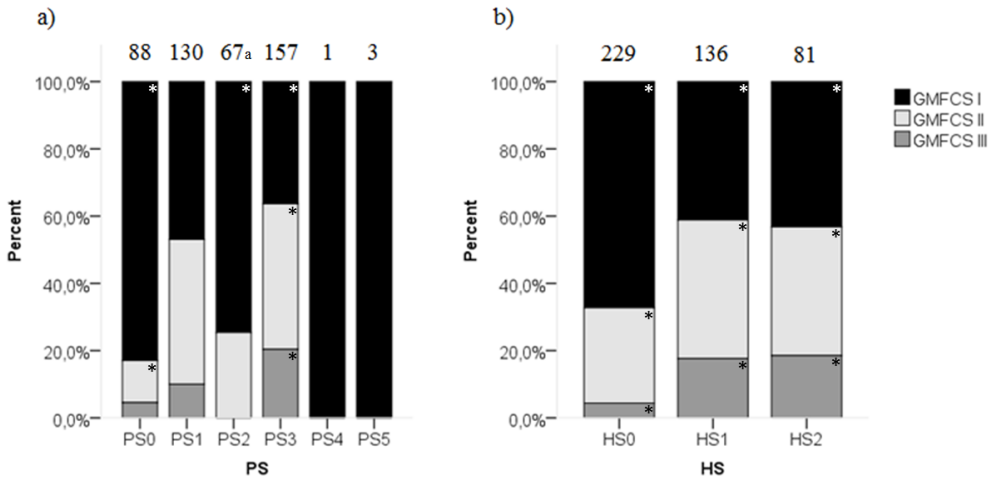


Figure 3. GMFCS level associated moderately with the distribution of **a)** pelvis patterns in sagittal plane (PS) and **b)** hip patterns in sagittal plane (HS). * indicates that a pattern was present significantly more or less frequently in patients with GMFCS level I, II, or III (p<0.05). [a] indicates that increased pelvic anterior tilt (PS2) was observed significantly less often in patients with GMFCS III. Specific ASRs are available in Table S2. Numbers on top of each bar represent the number of patients that were classified into that pattern.

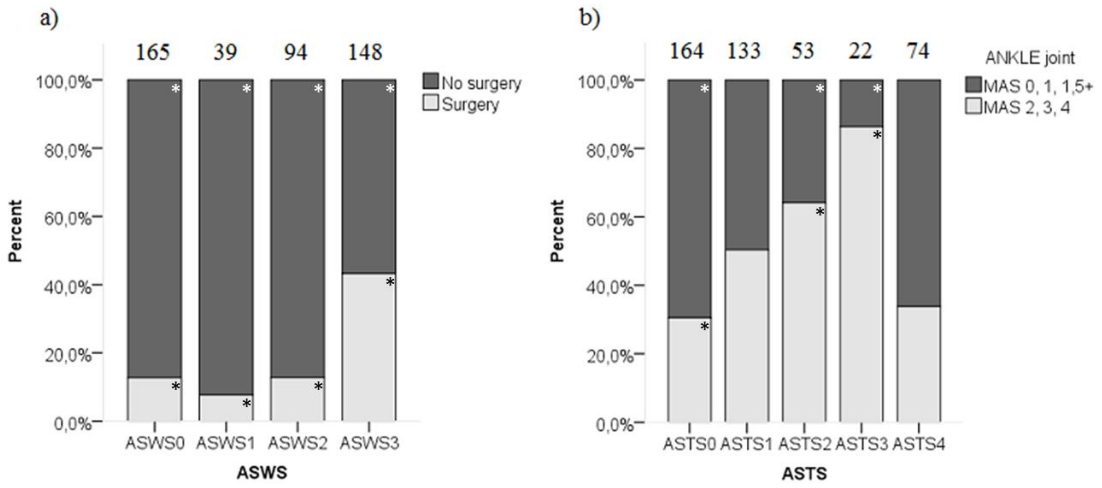


Figure 4. a) Previous surgery associated moderately with the distribution of the ankle patterns during swing (ASWS). **b)** Spasticity of muscles acting around the ankle associated moderately with the distribution of the ankle patterns during stance (ASTS). * indicates that a pattern was present significantly more or less frequently in limbs with or without surgery, or in limbs with lower (MAS 0, 1, 1+) vs. higher (MAS 2, 3, 4) levels of spasticity around the ankle ($p<0.05$). Specific ASRs are available in Table S4. Numbers on top of each bar represent the number of limbs that were classified into that pattern.

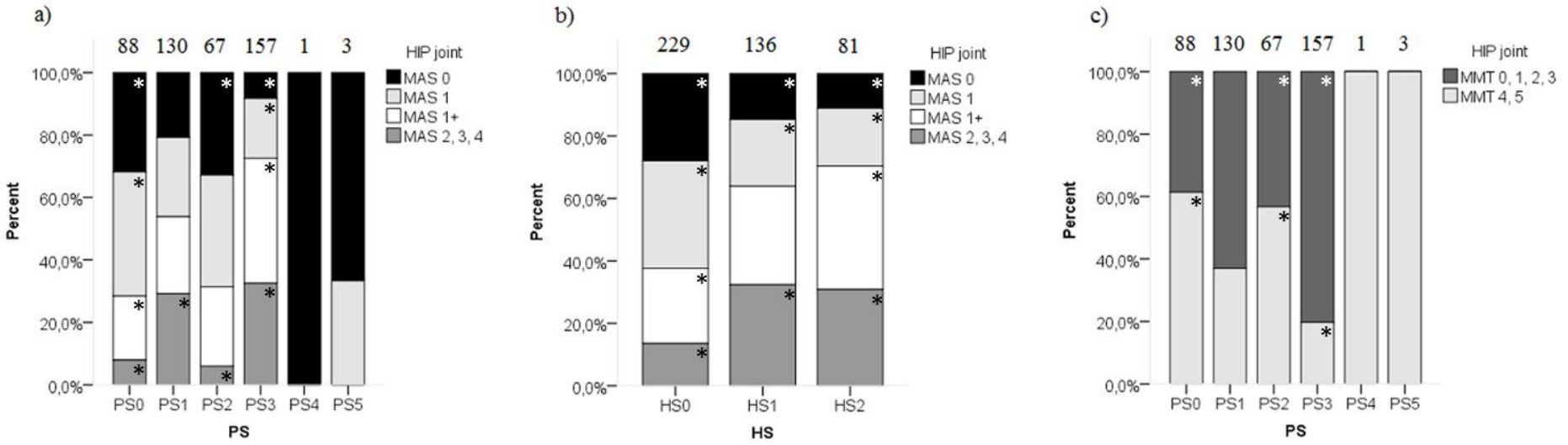


Figure 5. Spasticity of muscles acting around the hip associated moderately with the distribution of **a)** pelvis patterns in sagittal plane (PS), and **b)** hip patterns in sagittal plane (HS). **c)** Weakness of muscles acting around the hip associated moderately with PS. * indicates that a pattern was present significantly more or less frequently in limbs with of a particular MAS score or in limbs with weaker (MMT 0, 1, 2, 3) or stronger (MMT 4, 5) muscles around the hip ($p<0.05$). Specific ASRs are available in Table S2. Numbers on top of each bar represent the number of limbs that were classified into that pattern.

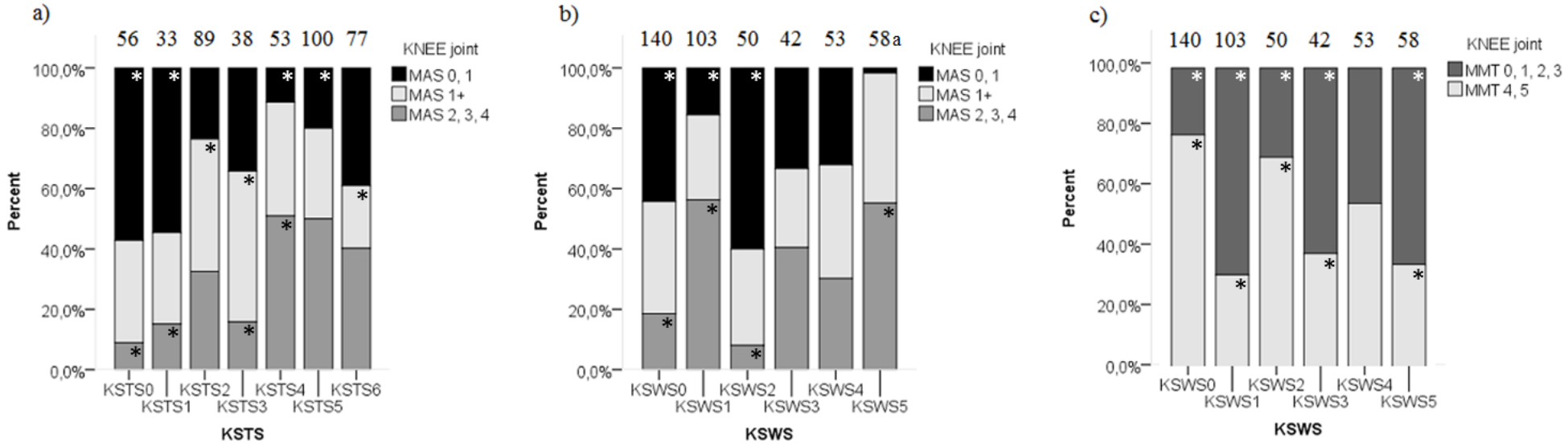


Figure 6. Spasticity of muscles acting around the knee associated moderately with the distribution of **a)** knee patterns during stance (KSTS), and **b)** knee patterns during swing (KSWS). **c)** Weakness of muscles acting around the knee associated moderately with KSWS. * indicates that a pattern was present significantly more or less frequently in limbs with of a particular MAS score or in limbs with weaker (MMT 0, 1, 2, 3) or stronger (MMT 4, 5) muscles around the knee ($p<0.05$). [a] indicates that decreased and delayed peak knee flexion (KSWS5) was observed significantly less often with limbs classified as MAS 0 or 1. Specific ASRs are available in Table S3. Numbers on top of each bar represent the number of limbs that were classified into that pattern.

Discussion

In this exploratory study, the prevalence of joint patterns during gait in children with CP and their association to patient-specific characteristics, previous surgery, and clinical symptoms, was examined.

The pattern ‘no or minor gait deviations’ was observed most frequently in all joints, apart from the knee during stance and pelvis patterns in the sagittal plane. The prevalence of ‘minor gait deviations’ reached more than 50% for the hip across the three anatomical planes, the pelvis in the coronal plane, and the foot progression angle. The need to define a pattern showing mild gait pathology has also been reported before, for example for the classifications of Winters et al.[22] (hemiplegic patterns) and Rodda et al.[23] (diplegic patterns)[24–26]. In both population- and hospital-based recruitment settings, the prevalence of these mild patterns has been reported to range between 12-43%[24,26]. The numbers in this study are generally higher, but this may be explained by the fact that the gait patterns in this study were evaluated at joint level, in contrast to the previously reported patterns at patient level which include multiple joints. In the present study however, a high number of ‘minor gait deviations’ in specific joints does not imply that most children with CP in this study walked closely to typical gait in general. Indeed, it was found that at patient level, only 6.7% of the included limbs were classified with ‘minor gait deviations’ in at least eight joints (out of eleven joints spread over the three anatomical planes), indicating that gait is markedly pathological in the majority of patients. So far, the way in which the various joint patterns across different planes combine in a total gait pattern is not yet fully understood.

Comparison of the prevalence of the pathological patterns to results from previous research is very challenging, as definitions of gait patterns as well as recruitment methods and inclusion criteria vary substantially across studies. For example, observed frequencies of excessive pelvic or hip rotation or in/outtoeing were markedly lower than reports by previous studies[12,27–29]. However, the definition of what constitutes excessive rotation across studies varies substantially. In the present study, a more strict definition was used by evaluating excessive rotation continuously over the entire gait cycle (or stance phase for FPA). This strict criterion is justified, taking into account the previously reported higher measurement errors for hip rotation and FPA[30]. A notable finding of the current study was that the pattern ‘decreased pelvic anterior tilt’ (or posterior tilt) with or without increased range of motion

was observed only four times. With these low numbers, the relevance of including both features as separate patterns in the classification could be questioned and should be re-examined. Posterior pelvic tilt was previously included as a potential feature of the type IV gait pattern defined by Rodda et al.[23], although it is unclear how often this feature is present in patients with type IV gait pattern[23,31]. The type IV pattern is mainly described for severely affected children. Following the assumption that posterior tilt will therefore be more prevalent in children with fewer functional abilities, the present study might have underestimated the prevalence of this pattern due to the relatively smaller sample size of children with GMFCS level III.

Relations with patient-specific characteristics and clinical symptoms

It was hypothesized that the prevalence of the patterns would be associated with age, topographical classification, and GMFCS level. This hypothesis could be confirmed for some joints, but the strength of most identified associations was weak. The knee patterns during swing and the pelvis patterns in the frontal and transverse plane showed moderate associations with topographical classification. Hence, they can be considered as characterizing for children with unilateral or bilateral CP. The finding that children with unilateral CP have a relatively higher prevalence of pelvic depression and excessive pelvic external rotation compared to children with bilateral CP concurs with previous research investigating hemiplegic gait[28,32,33]. The results further showed that the prevalence of the ankle patterns during stance associated moderately with age, with the youngest patients showing a relatively higher frequency of a horizontal or reversed second ankle rocker. Wren et al.[12] also noted decreased odds of equinus and increased odds of calcaneus gait with increasing age. The definition of equinus in their study (i.e. ankle plantarflexion >1 standard deviation below the mean for normal gait), would include the horizontal and reversed second ankle rocker, as well as the equinus pattern from the present study. These authors also reported an increased likelihood of presenting with internal hip rotation and/or outtoeing with increasing age[12]. The present study also found that intoeing occurred significantly less often than expected in older subjects, but no significant association was identified between hip patterns in the transverse plane and age. Different definitions of excessive internal hip rotation between both studies might again be the main cause of the marked differences in the observed frequency of this pattern (ca. 40% in Wren et al.[12], vs. 16.6% in this study). GMFCS levels are best characterized by the joint patterns in the sagittal plane. Although the results for the ankle and knee patterns should be interpreted with

caution, a trend showed that patterns with minor gait deviations at the level of each joint were mainly observed in children with GMFCS I.

The study also examined how specific gait patterns were characterized by weakness and spasticity. An obvious trend regarding all significant associations was that the patterns with minor gait deviations (PS0, HS0, KSTS0, KSWS0, ASTS0, ASWS0, PC0, PT0, HT0, FT0) were observed significantly more often in limbs with a low level of spasticity (MAS 0, 1, 1+) and good muscle strength (MMT 4 or 5), and significantly less often than expected in other pathological patterns. The pathological patterns that were most characterized by both weakness (MMT 0, 1, 2, or 3) and spasticity were patterns related to pelvic anterior tilt (PS2 and PS3), patterns with increased knee flexion at initial contact (KSTS1 and KSTS4), patterns with abnormal knee flexion in swing (KSWS1, KSWS2 and KSWS5), ankle patterns characterized by excessive plantar flexion (ASTS3 and ASWS2), and 'excessive hip internal rotation' (HT2). The patterns 'increased and delayed peak knee flexion during swing' (KSWS3) and 'outtoeing' (FPA1) were mainly characterized by weakness alone. On the other hand, 'reversed second ankle rocker' (ASTS2) and 'intoeing' (FPA2) were mainly characterized by spasticity. It was also apparent that stronger associations with clinical symptoms were consistently found for the joints in the sagittal plane, possibly because most of the evaluated muscles in this study also perform sagittal plane motions as a main function (i.e. flexion and extension around the hip, knee, and ankle).

Remarkably, there were no significant associations identified with any of the investigated variables for the hip in the coronal plane. A recent study evaluated the level of clinician agreement with which these patterns could be identified and found that the hip in the coronal plane had the highest number of 'unclassifiable' patients (paper under review in Developmental Medicine and Child Neurology). It has also been previously suggested that deviations in the coronal plane might be primarily characterized by compensatory movements for deviations in the sagittal or transverse plane, which are covered in other patterns[8]. Hence, the pattern definitions of the coronal plane patterns and their relevance or necessity in the classification should be re-examined.

Another hypothesis said that a specific joint would be associated in particular with the severity of weakness or spasticity in muscle groups that act around that joint. The results of this study confirmed that these associations were present, however, as table 4 and 5 demonstrate, joint patterns were also associated with weakness and spasticity scores of muscle

groups acting around the other joints. For instance, for the knee patterns during swing, a significant association was found with the level of spasticity for the muscles around the knee, but also with the level of spasticity around the ankle and hip joint. The directions of these significant associations were the same for the spasticity scores at each level: with higher scores of spasticity, the patterns 'delayed (and decreased)' peak knee flexion (KSWS1, KSWS5) were observed significantly more often; with lower scores of spasticity, the patterns 'minor gait deviations' (KSWS0) and 'increased peak knee flexion' (KSWS2) were more often observed. This finding can be extrapolated to all joint patterns: if joint patterns were associated with weakness or spasticity at more than one level (i.e. hip, knee, or ankle), the direction of the significant associations was similar for all levels (Tables S2-S6). This result suggests that specific gait deviations in one joint are not only caused by problems in the muscles surrounding that joint. They will rather be the result of a complex interplay of different muscles and movements at all lower limb joints.

Limitations

A few limitations of the study need to be addressed. The generalizability of the results of this study might be limited as the investigated study group was a sample of convenience, recruited from one hospital setting. Firstly, it was noted that there was an underrepresentation of patients with GMFCS III and an overrepresentation of patients with unilateral cerebral palsy in the studied sample compared to previously reported distributions of gross motor function and topographical classifications[34,35]. More clear trends with GMFCS level might be identified given a larger proportion of children with GMFCS III, especially for the knee patterns and for the ankle patterns during stance. Secondly, 70 of 356 patients were excluded, of which 14 patients (20%) were excluded due to missing data from the clinical examination. It was not possible to find out the precise reasons for these missing data (e.g. fatigue or age resulting in reduced collaboration of the child, oversight by clinician, etc.). As a result, a small bias towards the exclusion of weaker or more severely affected children in the studied sample cannot be excluded. Thirdly, because the study used retrospective data, a relatively large amount of patients had undergone previous Achilles tendon lengthening (29 out of 100 limbs that were operated upon). The generalizability of the results is therefore limited, as surgical strategies have evolved during the past ten to twenty years and tendon lengthening procedures are performed much less frequently[36,37]. It is therefore difficult to formulate strong conclusions regarding the influence of previous surgery on the distribution of the gait

patterns. In the future, the effect of previous surgery should be investigated using more specific subgroups regarding previous surgical interventions, or alternatively, prospective longitudinal intervention studies should be carried out to test the responsiveness of the patterns to different treatment interventions. In this study, it was further decided to group muscles at the level of each joint depending on their main function, and to select the most severe MAS or MMT score to represent the severity of spasticity or weakness at that joint. This implicates that when weakness at the level of the ankle is associated with specific ankle patterns, some of the scores used for statistical analysis might have been the result of ankle dorsiflexor weakness, others might have been due to ankle plantarflexor weakness. It is obvious that different muscles such as ankle plantar- and dorsiflexors would affect gait differently and potentially stronger associations might be discovered if these analyses would be performed on a muscle-specific rather than joint-specific basis. However, detailed investigations of the muscle-specific MMT and MAS scores around each joint revealed that problems of spasticity or weakness were mostly present in more than one muscle group. For example, the MMT score of the hip, knee, and ankle joint was defined based on the score of one muscle group only in 27.6%, 35.9%, and 25.8% all limbs. Only for the MAS score around the knee and the ankle joint, 78.3% and 54.0% of the scores were based on one muscle and over 90% of these particular scores were determined by spasticity in the hamstrings and gastrocnemius muscles. Yet, because several muscles are affected by weakness or spasticity to a similar extent, and because different categories of the MAS and MMT scale were merged, it can be assumed that muscle-specific analyses would not change the general interpretations of the currently presented results. Rather, they might point to specific muscles whose clinical characteristics are discriminating best between particular gait patterns. Lastly, the gait patterns for each limb were based on a single representative trial, whereas CP children are known to have a certain amount of variability across trials. Future research may evaluate to what extent this variability affects the classifications and how consistently these patterns are assigned across multiple trials.

Conclusion

The usefulness of any classification essentially relies on its potential to make distinctions between clinically relevant subgroups in CP. This study provided first insights toward the construct validity and clinical relevance of joint gait patterns in CP[9]. Although further validation is warranted, the results of this study confirm that most gait patterns are

characterized by different patient-specific characteristics and that they are often associated with gross categories of muscle weakness and spasticity.

Acknowledgements

The authors would like to thank Nele Hendrickx and Sofie Blomme for their assistance in collecting the clinical examination data that was used in this study. A. Nieuwenhuys is supported by an OT project of KU Leuven University (OT/12/100). E. Papageorgiou is supported by the MD Paedigree project, a Model-Driven Paediatric European Digital Repository, partially funded by the European Commission under FP7 – ICT Programme (grant agreement no: 600932, http://www.md-paedigree.eu) and by the SIMCP IWT-project, a simulation platform to predict gait performance following orthopedic intervention in children with cerebral palsy (IWT 140184). The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

1. Bax M, Goldstein M, Rosenbaum P, et al. Proposed definition and classification of cerebral palsy, April 2005. *Dev Med Child Neurol*. 2005;47(8):571-576.
2. Palisano R, Rosenbaum P, Walter S, Russell D, Wood E, Galuppi B. Development and reliability of a system to classify gross motor function in children with cerebral palsy. *Dev Med Child Neurol*. 1997;39(2):214-223.
3. Eliasson A-C, Krumlinde-Sundholm L, Rösblad B, et al. The Manual Ability Classification System (MACS) for children with cerebral palsy: scale development and evidence of validity and reliability. *Dev Med Child Neurol*. 2006;48(7):549-554.
4. Õunpuu S, Gorton G, Bagley A, et al. Variation in kinematic and spatiotemporal gait parameters by Gross Motor Function Classification System level in children and adolescents with cerebral palsy. *Dev Med Child Neurol*. 2015;57(10):955-962.
5. Dobson F, Morris ME, Baker R, Graham HK. Gait classification in children with cerebral palsy: A systematic review. *Gait Posture*. 2007;25(1):140-152.
6. Carriero A, Zavatsky A, Stebbins J, Theologis T, Shefelbine SJ. Determination of gait patterns in children with spastic diplegic cerebral palsy using principal components. *Gait Posture*. 2009;29(1):71-75.
7. Bonnefoy-Mazure A, Sagawa YJ, Lascombes P, De Coulon G, Armand S. Identification of gait patterns in individuals with cerebral palsy using multiple correspondence analysis. *Res Dev Disabil*. 2013;34(9):2684-2693
8. Davids JR, Bagley AM. Identification of Common Gait Disruption Patterns in Children With Cerebral Palsy. *J Am Acad Orthop Surg*. 2014;22(12):782-790.
9. Nieuwenhuys A, Õunpuu S, Van Campenhout A, et al. Identification of joint patterns during

gait in children with cerebral palsy: A Delphi consensus study. *Dev Med Child Neurol*. 2016;58(3):306-313.

10. Nieuwenhuys A, Papageorgiou E, Dimitriou G, Molenaers G, De Laet T, Desloovere K. Inter- and intrarater reliability of an international consensus-based gait classification system in children with cerebral palsy. *Gait Posture*. 2015;42:Supplement 1, S46.
11. Zwick EB, Leistritz L, Milleit B, et al. Classification of equinus in ambulatory children with cerebral palsy - Discrimination between dynamic tightness and fixed contracture. *Gait Posture*. 2004;20(3):273-279.
12. Wren TAL, Rethlefsen S, Kay RM. Prevalence of specific gait abnormalities in children with cerebral palsy: influence of cerebral palsy subtype, age, and previous surgery. *J Pediatr Orthop*. 2005;25(1):79-83.
13. Domagalska M, Szopa A, Syczewska M, Pietraszek S, Kidon Z, Onik G. The relationship between clinical measurements and gait analysis data in children with cerebral palsy. *Gait Posture*. 2013;38(4):1038-1043.
14. Rozumalski A, Schwartz MH. Crouch gait patterns defined using k-means cluster analysis are related to underlying clinical pathology. *Gait Posture*. 2009;30(2):155-160.
15. Dobson F, Morris ME, Baker R, Graham HK. Unilateral cerebral palsy: a population-based study of gait and motor function. *Dev Med Child Neurol*. 2011;53(5):429-435.
16. Graham HK, Harvey A, Rodda J, Nattrass GR, Pirpiris M. The Functional Mobility Scale (FMS). *J Pediatr Orthop*. 2004;24(5):514-520.
17. Cuthbert C, Goodheart J. On the reliability and validity of manual muscle testing: a literature review. *Chiropr Osteopat*. 2007;50(1):9-15.
18. Daniels L, Worthingham C. *Muscle Testing Techniques of Manual Examination*. Philadelphia, PA: W B Saunders; 1986.
19. Mutlu A, Livanelioglu A, Gunel MK. Reliability of Ashworth and Modified Ashworth scales in children with spastic cerebral palsy. *BMC Musculoskelet Disord*. 2008;9:44.
20. Portney L, Watkins M. *Foundations of Clinical Research: Applications to Practice*. Vol 2nd editio. Prentice Hall Health, Upper Saddle River, New Jersey; 2000.
21. Cohen J. *Statistical Power and Analysis for the Behavioral Sciences (2nd Ed.)*. Hillsdale, Erlbaum; 1988.
22. Winters TF, Gage JR, Hicks R. Gait patterns in spastic hemiplegia in children and young adults. *J Bone Joint Surg Am*. 1987;69:437-441.
23. Rodda JM, Graham HK, Carson L, Galea MP, Wolfe R. Sagittal gait patterns in spastic diplegia. *J Bone Joint Surg Br*. 2004;86(2):251-258.
24. McDowell BC, Kerr C, Kelly C, Salazar J, Cosgrove A. The validity of an existing gait classification system when applied to a representative population of children with hemiplegia. *Gait Posture*. 2008;28(3):442-447.
25. Riad J, Haglund-Akerlind Y, Miller F. Classification of Spastic Hemiplegic Cerebral Palsy in Children. *J Pediatr Orthop*. 2007;27(7):758-764.
26. de Morais Filho MC, Kawamura CMC, Lopes JJAF, Neves DLD, Cardoso MM de O, Caiafa JJB. Most frequent gait patterns in diplegic spastic cerebral palsy. *Acta Orthop Bras [online]*. 2014;22(4):197-201.
27. Simon A-L, Ilharreborde B, Megrot F, et al. A Descriptive Study of Lower Limb Torsional Kinematic Profiles in Children With Spastic Diplegia. *J Pediatr Orthop*. 2015;35(6):576-582.

28. O'Sullivan R, Walsh M, Jenkinson A, O'Brien T. Factors associated with pelvic retraction during gait in cerebral palsy. *Gait Posture*. 2007;25(3):425-431.

29. de Morais Filho MC, Kawamura CM, Andrade PH, Dos Santos MB, Pickel MR, Neto RB. Factors associated with pelvic asymmetry in transverse plane during gait in patients with cerebral palsy. *J Pediatr Orthop Part B*. 2009;18(6):320-324.

30. Schwartz MH, Trost JP, Wervey R a. Measurement and management of errors in quantitative gait data. *Gait Posture*. 2004;20(2):196-203.

31. Stott NS, Atherton WG, Mackey AH, Galley IJ, Nicol RO, Walsh SJ. The reliability and validity of assessment of sagittal plane deviations in children who have spastic diplegia. *Arch Phys Med Rehabil*. 2005;86(12):2337-2341.

32. Salazar-Torres JJ, McDowell BC, Kerr C, Cosgrove AP. Pelvic kinematics and their relationship to gait type in hemiplegic cerebral palsy. *Gait Posture*. 2011;33(4):620-624.

33. Graham HK, Baker R, Dobson F, Morris ME. Multilevel orthopaedic surgery in group IV spastic hemiplegia. *J Bone Jt Surg - Ser B*. 2005;87(4):548-555.

34. Gorter JW, Rosenbaum PL, Hanna SE, et al. Limb distribution, motor impairment, and functional classification of cerebral palsy. *Dev Med Child Neurol*. 2004;46:461-467.

35. Rosenbaum PL. Prognosis for Gross Motor Function in Cerebral Palsy: Creation of Motor Development Curves. *JAMA J Am Med Assoc*. 2002;288(11):1357-1363.

36. Gage JR. Con:Interobserver variability of gait analysis. *J Pediatr Orthop*. 2003;23(6):290-291.

37. Healy MT, Schwartz MH, Stout JL, Gage JR, Novacheck TF. Is simultaneous hamstring lengthening necessary when performing distal femoral extension osteotomy and patellar tendon advancement? *Gait Posture*. 2011;33(1):1-5.

Supporting Information

Table S1. Cut-off values to interpret the strength of a significant association between two variables using Cramer's V statistic are dependent on the degrees of freedom (DF) [21].

	Cramer's V	**Interpretation**
DF = 1	0.10 < V < 0.30	Weak association
	0.30 < V < 0.50	Moderate association
	V > 0.50	Strong association
DF = 2	0.07 < V < 0.21	Weak association
	0.21 < V < 0.35	Moderate association
	V > 0.35	Strong association
DF = 3	0.06 < V < 0.17	Weak association
	0.17 < V < 0.29	Moderate association
	V > 0.29	Strong association

DF is the smaller value of (R-1) or (C-1). R and C represent the number of categories of the related variables.

Tables S2-S6 indicate the direction of the significant associations between all gait patterns and the different categories of patient-specific characteristics (N=286), side-specific, and clinical variables (N=446). Numbers in the table represent the significant adjusted standardized residuals and indicate the specific combinations of categories and patterns that were observed more often (positive values) or less often (negative values) than would be expected if the variables were unrelated. Detailed cross-tables including the observed frequencies and percentages of the recruited sample population are available at request with the authors. Descriptions of the abbreviated patterns are available in the main paper in Table 1.

Table S2. Adjusted standardized residuals indicate the direction of significant associations between the sagittal plane patterns of the pelvis and hip joint and the categories of the patient-specific characteristics (N=286), previous surgery, and clinical variables (N=446).

		Pelvis[a]				Hip		
		PS0	**PS1**	**PS2**	**PS3**	**HS0**	**HS1**	**HS2**
N = 286	**Unilateral CP**					2.6		-2.5
Diagnosis	**Bilateral CP**					-2.6		2.5
Age	**youngest**	-2.1						-2.2
	medium aged			2.2				
	oldest	2.2		-2.6		-2.7	2.0	
GMFCS	**I**	4.2		2.4	-4.8	4.6	-3.3	-2.2
	II	-3.5			2.9	-2.2	2.0	0.5
	III			-2.3	3.5	-4.4	2.4	3.0
N = 446								
Surgery	**No**			2.6		2.6	-2.8	
	Yes			-2.6		-2.6	2.8	
MAS hip	**0**	3.0		2.8	-4.7	3.8	-2.1	-2.4
	1	2.9			-3.0	3.4	-2.0	-2.0
	1+	-2.1			3.7	-2.4		2.3
	2-3-4	-3.7	2.1	-3.5	3.7	-4.6	3.3	2.0
MAS knee	**0-1**	4.0		3.7	-4.3	3.3		-2.2
	1+							
	2-3-4	-3.3		-2.8	3.6	-5.1	3.1	2.9
MAS ankle	**0-1-1+**	2.8		3.6	-4.2			
	2-3-4	-2.8		-3.6	4.2			
MMT hip	**0-1-2-3**	-4.9		-3.3	6.1	-5.5	3.2	3.2
	4-5	4.9		3.3	-6.1	5.5	-3.2	-3.2
MMT knee	**0-1-2-3**	-4.7	3.5	-4.8	4.2	-5.2	5.7	
	4-5	4.7	-3.5	4.8	-4.2	5.2	-5.7	
MMT ankle	**0-1**	-2.1	2.4	-3.6	2.2	-3.7	3.0	
	2	-4.2			3.8	-3.5	2.4	
	3					2.1	-2.0	
	4-5	5.3		2.9	-5.7	4.1	-2.5	-2.4

MAS = modified ashworth scale score; MMT = manual muscle testing score; [a] N=282 patients and N=442 limbs due to exclusion of PS4 and PS5. Shaded area's indicate weak (light grey) and moderate (darker grey) associations based on Pearson Chi squared analyses.

Table S3. Adjusted standardized residuals indicate the direction of significant associations between the sagittal plane patterns of the knee joint and the categories of the patient-specific characteristics (N=286), previous surgery, and clinical variables (N=446).

		Knee during stance							Knee during swing					
		KSTS0	**KSTS1**	**KSTS2**	**KSTS3**	**KSTS4**	**KSTS5**	**KSTS6**	**KSWS0**	**KSWS1**	**KSWS2**	**KSWS3**	**KSWS4**	**KSWS5**
N = 286	**Unilateral CP**	2.3			2.9	-2.2	-3.2		3.9	-2.3				-3.4
Diagnosis	**Bilateral CP**	-2.3			-2.9	2.2	3.2		-3.9	2.3				3.4
Age	**youngest**								-2.6	4.0		2.9		
	medium aged									-2.3				2.6
	oldest								2.3					
GMFCS	**I**													
	II													
	III													
N = 446														
Surgery	**No**							-2.9						
	Yes							2.9						
MAS hip	**0**	4.3	2.3			-2.5	-2.2		4.5	-3.7	4.6			-3.5
	1		2.0				-2.2		3.1	-2.4				-3.1
	1+							2.1					2.1	
	2-3-4	-4.3	-2.8		-2.2	2.5	4.8		-6.7	5.6	-3.3		-2.4	6.4
MAS knee	**0-1**	4.4	3.0			-3.4	-2.8		4.0	-4.0	4.6			-5.2
	1+			2.1	2.1			-2.7						
	2-3-4	-4.3	-2.4		-2.5	2.7	3.8		-4.7	5.4	-4.2			3.6
MAS ankle	**0-1-1+**	4.2	3.8			-2.9			5.2	-5.4	3.6			-4.2
	2-3-4	-4.2	-3.8			2.9			-5.2	5.4	-3.6			4.2
MMT hip	**0-1-2-3**	-4.7	-3.7			3.2	2.8		-5.0	4.0	-3.2	2.2		3.1
	4-5	4.7	3.7			-3.2	-2.8		5.0	-4.0	3.2	-2.2		-3.1
MMT knee	**0-1-2-3**	-3.9	-2.2			3.1	2.7		-6.8	5.3	-2.4	2.2		3.2
	4-5	3.9	2.2			-3.1	-2.7		6.8	-5.3	2.4	-2.2		-3.2
MMT ankle	**0-1**	-2.4				2.1			-3.8	2.5				3.2
	2	-2.8				2.9			-2.5		-2.9			
	3													
	4-5	4.7	3.3			-3.1	-2.0		5.1	-4.6				-3.0

MAS = modified ashworth scale score; MMT = manual muscle testing score; Shaded area's indicate weak (light grey), moderate (darker grey), and strong (darkest grey) associations based on Pearson Chi squared analyses.

Table S4. Adjusted standardized residuals indicate the direction of significant associations between the sagittal plane patterns of the ankle joint and the categories of the patient-specific characteristics (N=286), previous surgery, and clinical variables (N=446).

		Ankle during stance					Ankle during swing			
		ASTS0	**ASTS1**	**ASTS2**	**ASTS3**	**ASTS4**	**ASWS0**	**ASWS1**	**ASWS2**	**ASWS3**
N = 286	**Unilateral CP**							2.9	2.8	-2.7
Diagnosis	**Bilateral CP**							-2.9	-2.8	2.7
Age	**youngest**	-2.0	2.1	3.1		-2.8				
	medium aged									
	oldest	2.6	-2.0	-2.2						
GMFCS	**I**									
	II									
	III									
N = 446										
Surgery	**No**		2.7			-3.5	3.8	2.3	2.5	-7.4
	Yes		-2.7			3.5	-3.8	-2.3	-2.5	7.4
MAS hip	**0**	4.5		-2.9						
	1									
	1+			2.1						
	2-3-4	-3.7			2.7					
MAS knee	**0-1**	3.1	-2.8		-2.3	2.1	3.2		-3.1	
	1+								2.6	
	2-3-4	-3.6	2.5		2.5		-2.2			2.0
MAS ankle	**0-1-1+**	4.3		-3.2	-4.1		2.2		-4.0	
	2-3-4	-4.3		3.2	4.1		-2.2		4.0	
MMT hip	**0-1-2-3**						-3.1			3.1
	4-5						3.1			-3.1
MMT knee	**0-1-2-3**	-3.4			2.6	2.0	-2.5			3.0
	4-5	3.4			-2.6	-2.0	2.5			-3.0
MMT ankle	**0-1**	-2.9					-3.4		2.1	2.1
	2				2.7		-2.4		2.1	
	3									
	4-5	3.0			-2.6		4.1		-2.5	

MAS = modified ashworth scale score; MMT = manual muscle testing score; Shaded area's indicate weak (light grey) and moderate (darker grey) associations based on Pearson Chi squared analyses.

Table S5. Adjusted standardized residuals indicate the direction of significant associations between the coronal plane patterns and the categories of the patient-specific characteristics (N=286), previous surgery, and clinical variables (N=446).

		Pelvis				Hip			
		PC0	**PC1**	**PC2**	**PC3**	**HC0**	**HC1**	**HC2**	**HC3**
N = 286	**Unilateral CP**				4.8				
Diagnosis	**Bilateral CP**				-4.8				
Age	**youngest**								
	medium aged	-2.2			2.1				
	oldest	2.3							
GMFCS	**I**								
	II								
	III								
N = 446									
Surgery	**No**	-2.4			2.0				
	Yes	2.4			-2.0				
MAS hip	**0**		-2.8		2.6				
	1								
	1+	-2.8	3.3						
	2-3-4								
MAS knee	**0-1**	3.3	-3.4						
	1+	-2.4							
	2-3-4		2.1		-2.1				
MAS ankle	**0-1-1+**								
	2-3-4								
MMT hip	**0-1-2-3**	-2.3	2.7						
	4-5	2.3	-2.7						
MMT knee	**0-1-2-3**			2.6					
	4-5			-2.6					
MMT ankle	**0-1**								
	2								
	3								
	4-5								

MAS = modified ashworth scale score; MMT = manual muscle testing score; Shaded area's indicate weak (light grey) and moderate (darker grey) associations based on Pearson Chi squared analyses.

Table S6. Adjusted standardized residuals indicate the direction of significant associations between the transverse plane patterns and the categories of the patient-specific characteristics (N=286), previous surgery, and clinical variables (N=446).

		Pelvis				Hip			Foot progression angle		
		PT0	**PT1**	**PT2**	**PT3**	**HT0**	**HT1**	**HT2**	**FT0**	**FT1**	**FT2**
N = 286	**Unilateral CP**		-2.8	3.9	-2.9				2.1		-3.8
Diagnosis	**Bilateral CP**		2.8	-3.9	2.9				-2.1		3.8
Age	**youngest**										2.7
	medium aged										
	oldest										-2.6
GMFCS	**I**										
	II										
	III					-3.4	2.0	2.6			
N = 446											
Surgery	**No**					2.1	-3.2				
	Yes					-2.1	3.2				
MAS hip	**0**					3.4		-3.9			-3
	1								2.0		-2.1
	1+					-2.3		3.2			2.5
	2-3-4					-2.9		2.3	-2.7		2.5
MAS knee	**0-1**					2.1		-2.8			-2.6
	1+	2.0	-2.7								
	2-3-4	-3.6	2.7			-3.2		3.7	-2.0		2.9
MAS ankle	**0-1-1+**					2.4		-3			
	2-3-4					-2.4		3			
MMT hip	**0-1-2-3**	-3.1	2.0			-3		2.6			
	4-5	3.1	-2.0			3		-2.6			
MMT knee	**0-1-2-3**					-4.1		3.3	-2.4	2.4	
	4-5					4.1		-3.3	2.4	-2.4	
MMT ankle	**0-1**			2.6		-4		3.4	-2.0	2.4	
	2	-2.9		2.5							
	3					2.6					
	4-5	3.1		-2.8				-2.3	2.4		

MAS = modified ashworth scale score; MMT = manual muscle testing score; Shaded area's indicate weak (light grey) and moderate (darker grey) associations based on Pearson Chi squared analyses.

Chapter 7

General discussion

Synopsis

The heterogenic clinical presentation of CP is especially striking when looking at gait. The golden standard to evaluate gait in children with cerebral palsy (CP) is three-dimensional gait analysis (3DGA). The principal goal of this PhD thesis was to develop a clinically relevant, valid, and reliable classification system for pathological movement patterns during gait in children with spastic CP, based on 3DGA data. The possible applications and advantages of a standardized classification are plenty. Among others, it supports medical practitioners in their clinical reasoning and eases communication among health care workers, patients, and their families[1,2]. It also presents a uniform terminology, which is useful for research purposes and can facilitate a more transparent interpretation of scientific literature[3].

The trigger for this thesis was the variability in, and difficulty of interpreting and analyzing the large amount of 3DGA data, which is characterized by multidimensionality as well as non-linear and time-dependent relations between waveforms[4,5]. Typically, 3DGA data is reduced through the analysis of gait features and gait classifications. Several problems that are related with these reduction methods were recognized (see also Chapter 1, Introduction):

(1) There is no consensus among clinicians or researchers on the definitions of kinematic or kinetic gait features and patterns that are responsive to change after treatment or that are clinically relevant to characterize CP gait.
(2) Feature analysis is biased by subjective feature selection. The increased probability of false positive outcomes, which is associated with the analysis of multiple dependent gait features, is difficult to control without decreasing statistical power.
(3) The scope of gait classifications in CP seems to range from patterns that are based on one specific gait feature to patterns that involve multiple gait features across several joints in more than one plane.
(4) Classifications developed via quantitative methods are criticized for identifying gait classes of which are not interpretable by clinicians. Qualitative classifications are criticized for being subjective and vague during the classification development process.
(5) Psychometric properties of gait classifications are usually not examined. In addition, several methodological shortcomings were identified for studies that examined content validity, construct validity, or responsiveness of a classification.

Within this thesis, there are five studies that contributed to the development of a new gait classification for ambulatory children with spastic CP. By profiting from the combined potential of sound qualitative and quantitative research methods, the abovementioned problems related to the reduction, analysis, and interpretation of 3DGA data were accounted for.

The first study presented a comprehensive overview of clinically relevant gait features and examined an alternative, quantitative method to extract clinically relevant information from kinematic and kinetic waveforms[6] (part 1). Incorporating knowledge gained from literature reviews and local expert meetings[6,7], the second study described the development process of a new qualitative gait classification[8] (part 2). The final three studies explored the reliability, content validity, and construct validity of the developed classification (part 3). This general discussion starts by summarizing the results for each study. The elements separating the research studies from previously published investigations are presented, but are also critically reflected upon. The second part of the discussion considers general methodological limitations, after which future research steps are proposed and an overall conclusion is formulated.

Summary and critical reflections

Part 1 - Gait features

In literature, gait features are commonly extracted from 3DGA data on a subjective basis to determine changes in gait after treatment. This first part of the PhD project searched for a more objective, alternative approach to extract clinically relevant information form kinematic and kinetic waveforms[6]. In a first step, a systematic literature search was performed to create an overview of all gait features that had previously been reported in literature, and that were demonstrated to be responsive to Botulinum Toxin type A (BTX-A) treatment. Twenty-six intervention studies evaluating the effect of BTX-A treatment on gait using 3DGA were identified, and within those papers, 53 kinematic and 33 kinetic features were reported. On several occasions, features were not adequately defined. In addition, the gait features that were selected for statistical analysis by each of those 26 studies, varied substantially, even though they mostly evaluated the effect of BTX-A injections in the same muscles (i.e. triceps surae). In a second step, a retrospective intervention study was carried out, comparing the results of two statistical approaches to analyze the effect of BTX-A injections on gait using

3DGA. Regarding the first approach, all gait features that were previously identified by the literature review, were evaluated by paired samples t-tests, using the Holm's correction to correct for an increased probability of false positives[9]. For the second approach, statistical parametric mapping (SPM) was used to analyze the full kinematic waveforms, thereby accounting for the dependency of each point of the waveform (i.e. controlling the probability of obtaining false positive results) and avoiding the potential bias of a priori data reduction by selecting features[10]. The study found that both approaches largely lead to similar conclusions and both feature analysis and SPM analysis can thus be valid approaches if they are commensurate to an a priori stated hypothesis. However, SPM analysis also identified a few additional, significant areas during the gait cycle at the level of the knee, which were not previously reported in any of the studies identified during the literature review.

Reflections

In a parallel study (reported as a master thesis in Rehabilitation Sciences and Physiotherapy at KU Leuven), the methodology of this first part of the PhD project was also applied to examine the effect of selective dorsal rhizotomy on gait in children with CP. This study drew similar conclusions, namely that feature definitions are not always clear, and that SPM identifies additional clinically relevant information from the kinematic and kinetic curves that were not included before in features extracted from literature.

The difference between these two studies and previous literature is that feature selection for these studies was based on an extended systematic literature search. It can therefore be assumed that feature selection was less biased because of the combination of all available expert knowledge from the different studies. This is in contrast to the 26 previously published studies on BTX-A treatment, for which the median number per study was only five features, which were (presumably) subjectively selected. Only five studies stated that feature selection was based, in part, on previous findings in literature[11–15]. It is therefore possible that previous studies might have slightly under- or overestimated gait changes due to treatment, depending on the features that were chosen and whether or not an increased probability of false positives rates was taken into account. This was also recently shown by Pataky et al.[16], who quantified the probability of false positive rates when using zero-dimensional feature analysis for hypotheses that pertain to one-dimensional (i.e. time-varying) waveforms. To conclude, SPM is a valid and more objective approach to analyze the effect of treatment on gait, especially when clinical expert knowledge on gait biomechanics in a particular patient population is

limited. In contrast to many data reduction techniques (such as principal component analysis or summary indices (e.g. Gait Profile Score[17])), results emerging from SPM are also interpreted directly in relation to the kinematic and kinetic waveforms, which makes it an attractive tool for clinicians.

Two shortcomings, which are relevant for both feature analysis and SPM analysis, need to be addressed. First, results of statistical feature analysis describe the probabilistic behavior of random data (i.e. p-value). Results of SPM analysis also describe the probability of detecting a cluster with a specific temporal length from random data, and uses random field theory to account for the interdependency of each point of a one-dimensional waveform[18]. However, there is no theory that describes how a musculoskeletal system with multiple interrelated joints will randomly behave. SPM can model the interdependency of vectors, such as the pelvis kinematics in the sagittal, coronal, and transverse plane, yet it cannot take into account the interdependency between joints[19,20]. In analyzing a large number of interdependent joints, statistical power is threatened once more, especially when analyzing relatively small experimental populations[21]. This effect could also have played with the presented study.

The second remark concerning both approaches is that statistical significance is not equal to clinical relevance per se. While minimal clinically important differences have been reported for clinical outcome tools such as the Gross Motor Function Measure[22], and for summary measures based on 3DGA such as the Gait Profile Score[17], they have not been commonly quantified for the analysis of gait features, or for SPM analysis. Klejman et al.[23] have quantified minimal detectable changes for 25 kinematic features in a group of 28 children with spastic CP. They found an average minimal detectable change of 10.5° (range 3.9°-16.1°). Sutherland et al.[24] reported that they considered changes to be clinically relevant if the change in joint angles was larger than 3°, because this was equal to the maximum inter-session error measured at their laboratory. These are good solutions if the standard inter-session and inter-therapist measurement error of a gait laboratory has been quantified. From previous research by McGinley et al.[25] and Kaufman et al.[26], it can be concluded that most studies report standard measurement errors for 3DGA waveforms of less than 5°, except for hip and knee rotation, the latter not being considered within this dissertation. An assessment of clinical relevance was not performed for the study in Chapter 2. However, for the outcome of SPM analyses in Chapter 5, an average difference of at least 3° within the areas of significance, highlighted by the SPM output, was judged visually as a threshold to determine

clinical relevance. Considering literature findings, a difference of 3° might be the result of measurement error on some occasions, yet on the other hand, a change in joint motion of even 2° could be clinically very relevant for a patient as it might represent the difference between foot clearance and toe drag during swing. The threshold of 3° was therefore chosen as a compromise between plausible measurement error and the potential clinical relevance of small changes in gait.

Part 2 - Classification development process

The second part of the PhD project concerned the development of a new gait classification in children with CP. By means of a Delphi consensus study, an international expert panel defined joint patterns during gait that were deemed to be clinically relevant and characteristic of gait pathology in children with CP[8]. Delphi studies employ iterative surveys to measure consensus on a given topic, and with each survey, participants receive the results of the previous round[27]. This allows each member to reflect on their own opinion in light of the opinion of other panel members. Before the start of the Delphi study, a preliminary proposal of joint patterns during gait was defined by clinical experts of university hospital Pellenberg. On top of the experience of the clinical team, the step-by-step process that was undertaken to define this preliminary proposal, was founded on previously published classifications and terminology, as well as on gait waveforms and gait features of CP and typically developing (TD) children[7]. At the end of the Delphi study, gait patterns were defined at the level of each joint in the three anatomical planes. Three to seven patterns were defined per joint and all patterns reached at least 75% agreement, apart from one of the knee patterns during stance.

This Delphi consensus project that was organized to develop a new gait classification in CP is distinct from previous qualitatively developed classifications since it was the first study to transparently describe the development process. Unlike other qualitative classifications, it was clear which gait deviations were chosen, how they were chosen, and who chose them[2,28–35]. Another unique aspect was that the study was a consensus study, combining clinical expert knowledge from eight different gait laboratories across the USA and Europe. Furthermore, the classification was intended for all ambulatory children with spastic CP and pattern definitions included all lower limb joints and deviations across the three anatomical planes, which was not reported before in a qualitative classification, except for the recently described classification by Davids and Bagley[28]. The Delphi consensus approach had not previously been used to develop gait classifications, yet it is not new in the field of CP research. It was

also used in the past years to define classifications of eating and drinking ability as well as communicative function classification systems[36–38]. In fact, one of the most cited classifications in CP research, the GMFCS, was also based on nominal group processes and Delphi surveys[39,40]. In all these studies, consensus approaches have formed a crucial stepping stone for further research that can establish the psychometric properties of a new classification.

Reflections

The patterns of the Delphi study should be considered as categorical patterns on a nominal scale, similar for example to the knee patterns defined by Sutherland et al.[35]. Such nominal patterns at the level of each joint incorporate pathological deviations in any direction, but this also implies that severity is not taken into account, as opposed to the classification for instance by Winters et al. where children with a Type I pattern are less severely affected (i.e. distal joint deviations) than children with a Type IV gait pattern (i.e. proximal and distal joint deviations)[29]. Consequently, care should be taken toward the choice of statistical analysis methods when using the classification in research, as options are often more limited (less powerful) for nominal scales[41].

At the beginning of the project, the interpretability and clinical relevance of the classification that would be developed was deemed a priority. Therefore a qualitative development process was preferred over quantitative techniques[2]. To ensure a documented and transparent reporting of the development process, two consensus approaches were considered: the Delphi process and the nominal group technique[42]. Both approaches rely on an expert panel, but where the Delphi process uses iterative anonymous surveys to measure consensus, the nominal group technique uses structured meetings[43,44]. Within this PhD research, the Delphi approach was preferred over the nominal group technique because it minimizes the effect of peer pressure among panel members, and because it was more feasible from a practical point of view, as it does not require experts to be in the same place and at the same time on several occasions. A disadvantage of the Delphi process is that questions might be misinterpreted. To minimize this potential source of bias, a few actions were undertaken based on previously reported guidelines[27,45]. After each question, experts were asked to provide written comments if questions were unclear, or if they had any other suggestions. Furthermore, two students of the Master program in pediatric physical therapy at KU Leuven received a pilot version of each survey to identify unclear phrasing. In addition, before the start of the Delphi, the aims

of the study were presented to all panel members, as well as a detailed presentation on the preliminary classification, which was illustrated with kinematic and kinetic examples for each of the patterns.

For the Delphi study, the subjectivity of the approach, allowing experts to indicate the clinically relevant patterns in CP, is as much an asset as it is a liability. Members of the expert panel were all highly experienced and respected international researchers and teachers. All participants have been involved, be it alongside each other or not, in international gait analysis courses for several years, which might have sped up the consensus process. It will likely have caused them to develop a common language, aiding consensus on the specific terminology that was adopted for the definitions of the gait patterns. On the other hand, the Delphi remains above all dependent on the composition and size of the expert panel[27,45], an issue which was also discussed in the publication of the Delphi study[8], and underlined in a commentary by Chambers[3]. Due to this subjectivity, and in combination with the knowledge that no patient data was used to support the existence of the patterns in children with CP, the content validity of the classification remained unclear.

Part 3 - Psychometric properties

Three studies were set up to examine the reliability, content validity, and construct validity of the developed classification, which is in contrast with the majority of previously reported classifications, for which a combination of psychometric properties have often not been examined (cfr. Chapter 1; Introduction)[2]. First, the summaries and highlights of these studies are introduced. Then, critical reflections will be formulated based on the combined results of all three studies.

Reliability

An international agreement study quantified the level of inter- and intrarater agreement with which clinicians could assign the gait patterns of the Delphi study, using 3DGA data. An experimental group of 82 patients with CP was recruited and 32 clinical raters were asked to classify a subset of 27 or 28 patients twice, using a custom-made online graphical user interface (www.cmal-tools-leuven.be). Inter- and intrarater agreement was good to excellent for all joints, except for the knee during stance phase, for which the interrater agreement level was moderate. Results were similar for the mean interrater agreement between each rater and a criterion (expert) classification, except for the agreement of the pelvis patterns in the sagittal plane, which was found to be borderline moderate instead of good. Raters who were

experienced with the interpretation of 3DGA data performed better on the patterns of the knee during stance and of the ankle during swing, and performed similarly for all other joint patterns. Even though the results showed overall good agreement, poor interrater agreement levels were found for some specific patterns, mostly at the level of the knee during stance.

Strong elements in the design of this study were firstly, that a large international group of clinicians was recruited, who had not been involved in the development process of the classification. Secondly, the raters had varying levels of experience with the collection or interpretation of 3DGA data, and with the evaluation or treatment of children with CP. Thirdly, the study had sufficient power to allow for precise (more confident) interpretations of the results, as opposed to for instance the reliability studies on the classification of Winters et al.[29] and Rodda et al.[31], for which wide confidence intervals surrounding the agreement estimates were previously reported[46,47]. In combination with a brief online learning phase, these elements all maximize the generalizability of the results of this reliability study.

It is especially interesting that raters with lower levels of experience with 3DGA performed equally well for most joints in comparison to the more experienced group. This is however not uncommon in literature. Even though experienced raters might be hypothesized to do better, it has been suggested that they could be internally biased and are more prone to develop individual interpretations of a classification[46]. The clinical raters participating in the reliability study found a small number of trials to be unclassifiable (approximately 5%), which is a considerably lower amount than previously reported numbers, for instance on the classification of Winters et al.[29 46,48,49]. This low number of unclassifiable trials was promising for the study that assessed the content validity of the classification.

Content validity

Even though the classification included patterns for the different lower limb joints in the three anatomical planes, the content validity still needed to be examined because of the subjectivity of Delphi study, which was also conducted without the use of patient data (Cfr. Classification development, supra)[8]. To investigate the content validity, two experienced raters classified 1719 kinematic and kinetic trials of 356 children with CP. At the level of each joint, trials were classified as ‘no or minor gait deviations’ if they did not meet the criteria of any other pathological pattern. Afterwards, the mean kinematic and kinetic waveforms for each pattern and the pattern of TD children were analyzed using SPM to verify (1) whether the existence of the patterns and the subjective rules, which were defined during the consensus study, could

be confirmed and (2) whether potential patterns and relevant information might have been missed. The results indicated that for each pattern, all key locations that were included in the pattern definitions, were also indicated as significant areas by the SPM analysis. Nonetheless, additional locations, which were not included in the pattern definitions, were also highlighted by SPM. Suggestions to further refine definitions for the patterns of the knee during stance and during swing were discussed. Given the abovementioned difficulty in determining a threshold for clinical relevance (cfr. General Discussion, Part 1), refinement of pattern definitions were only advised in those cases where a clinician would be aided in discriminating one particular pattern from the other patterns. In this way, the reliability of using the classification could potentially be increased, while maintaining its simplicity and interpretability.

Apart from Morais Filho et al.[50], who examined the patterns of Sutherland et al.[35], this is one of a few studies that has evaluated the content validity of a qualitatively developed classification using statistical analyses on such a large, heterogeneous sample of objective patient data. The retrospective patient trials that were included in the study of this PhD research were collected for a variety of reasons: re-evaluations, pre- or post-single event multilevel surgery, pre- or post-BTX-A treatment, or pre- or post-selective dorsal rhizotomy. This suggests that the wide range of potential deviations in children with spastic CP was likely included in this study, adding to the generalizability of the obtained outcome. Riad et al.[49] have previously provided objective evidence based on kinematic and kinetic gait features to demonstrate objective differences between the gait patterns of the classification of Winters et al.[29]. Other authors examining the validity of gait classifications in CP focused more on the construct validity of the classifications[48,50–53].

Construct validity

The final study explored the construct validity of the classification by examining the associations between consensus-based gait patterns on the one hand, and patient-specific characteristics and clinical symptoms on the other hand. The prevalence of the gait patterns was evaluated in a retrospective sample of 286 patients with CP. The patterns with ‘no or minor gait deviations’ were most frequently observed at the level of each joint, except for the knee during stance phase and pelvis in the sagittal plane. The prevalence of the patterns ‘decreased pelvic anterior tilt’, or ‘decreased pelvic anterior tilt and increased range of motion’ was too low to be included in the analysis. Subsequently, the distribution of the gait

patterns at the level of most joints was found to be significantly associated with topographical classification, GMFCS level, age, previous orthopedic surgery, and the severity of spasticity and muscle weakness of muscles around the hip, knee, and ankle joints. The strongest associations were found between the patterns of the joints in the sagittal plane, and the scores of spasticity and muscle weakness. Only the hip patterns in the coronal plane did not associate significantly with any of the investigated variables.

Similar to the study evaluating the content validity of the classification, this study was able to perform analyses on a heterogeneous experimental population, including patients that were evaluated pre- or post-treatment, thus maximizing generalizability. In a more ideal situation, construct validity would be better evaluated using more objective, instrumented measures of spasticity and muscle weakness[54]. However, currently available measurement systems have not been implemented in the routine clinical examination protocol alongside a 3DGA session and would therefore not have allowed for a large patient population to be evaluated. Through the detailed interpretation of the results using adjusted standardized residuals, specific patterns were identified, which are better able to distinguish between weaker and stronger children, younger and older children, etc. The results of this study can therefore be used to develop new hypotheses for future research in order to further demonstrate the clinical applicability or responsiveness of the classification.

Reflections

Patient data that were used to evaluate the psychometric properties in Chapters 4-6 were always classified on a trial-by-trial basis. However, pathological gait of patients with CP is characterized by inter-trial (i.e. intra-subject) variability[55,56]. Redekop et al.[57] concluded that in children with CP, four to six strides should be averaged when analyzing gait features from 3DGA, to obtain a reliable estimate of those features that is representative of a patient's gait pattern. The gait patterns of the developed classification sometimes rely on one specific feature, but more often they are a combination of multiple features, and/or take into account the shape and position of the entire kinematic waveform. It is unclear to what extent a patient's inconsistency (i.e. intra-subject variability) might cause the patient to be classified into different patterns across multiple trials. Therefore, it is also not known how this intra-subject variability might have affected the results of the studies described in Chapters 4-6. Future research should quantify the inter-trial reliability of the classification system.

In addition to the classification on a trial-by-trial basis, Chapters 4-6 classified each patient trial into one distinct gait pattern. This is not an easy task as kinematic and kinetic data have a continuous distribution, and differences, such as for example the contrast between 'normal' or 'increased' range of motion, are difficult to judge visually. While some patients are textbook examples of a particular pattern, others will be borderline and might appear to fall somewhere in-between, an issue which was previously highlighted by several authors[3,58,59]. This requirement of a 'hard' assignment may explain a substantial part of the disagreement found in the reliability study. During the reliability study, raters were able to indicate on a five-point ordinal scale how confident they felt about their class allocation for each patient. In case they did not feel confident, they were able to indicate the pattern which they felt was also a likely possibility. Unfortunately, this data was not provided by every rater so the extent to which 'borderline cases' have determined the level of clinician agreement in that study could not be estimated with certainty. Instead, a conservative, 'hard' assignment was maintained during the analysis. The level of clinician agreement could be expected to improve if a factor of uncertainty is taken into account and a more 'soft' classification is allowed.

Apart from rater disagreement due to the continuous distribution of kinematic data, disagreement could have also arisen for example if pattern definitions were misinterpreted, if they were too vague, or if raters made accidental errors using the online graphical user interface. To distinguish between these sources of disagreements, confusion matrices can be examined, which portray the agreement between each rater and the criterion classification for each patient. If agreement is perfect (100%), all ratings within the confusion matrix would be contained in the diagonal cells from the upper left to lower right corner (for example HC0xHC0, HC1xHC1, etc. in Table 1). Table 1 presents an example of the confusion matrix of the hip in the coronal plane based on the results from the reliability study (Chapter 4). Disagreement on the patterns continuous excessive hip abduction and adduction are considered to be accidental because the kinematic differences between these patterns are so large (i.e. they represent opposite movements and lie on different sides of the one standard deviation reference band of TD children). From table 1, it is therefore clear that few accidental errors occurred (cells in light grey). It is also apparent that most disagreements arose from patients that were classified as having no or minor gait deviations by the experts (cells in dark grey). This gives an indication of the difficulty of defining what constitutes a "continuous" deviation. The reliability study (Chapter 4) reported 13% of all ratings of the hip in the coronal plane to be unclassifiable. In most cases, raters felt that excessive abduction or

adduction occurring solely during stance phase could not to be categorized as a 'minor deviation'. It is possible that several raters – instead of indicating a patient as unclassifiable – still indicated the "continuous" patterns if deviations only occurred during stance (Table 1, dark grey cells). In addition, in Chapter 6, it was reported that the distribution of the hip patterns in the frontal plane were not associated with any of the investigated variables, which questioned their construct validity and clinical relevance. Davids and Bagley[28] also considered the hip in the coronal plane to be of less relevance for clinical decision making.

Table 1. Confusion matrix of the hip patterns in the coronal plane visualizes the confusion of the rater group (experienced (n=15) and inexperienced (n=14) raters) with respect to the criterion classification, based on the results from Chapter 4. For instance the cell in column HC3 and row HC2 indicates that 4 four trials were classified as 'continuous excessive hip adduction' (HC3) according to the criterion classification, while inexperienced or experienced clinical raters indicated they belonged to the pattern 'continuous excessive hip abduction' (HC2). Cells in dark grey indicate most commonly observed disagreements; cells in light grey indicate likely accidental disagreements.

		Criterion classification			
		HC0	HC1	HC2	HC3
Rater group (n=29)	HC0	396	17	3	0
	HC1	50	75	6	0
	HC2	28	20	65	4
	HC3	46	2	4	74

HC0 = no or minor gait deviations
HC1 = excessive abduction during swing
HC2 = continuous excessive hip abduction
HC3 = continuous excessive hip adduction

Although good content validity of the classification was expected and generally confirmed, SPM analysis in Chapter 5 still showed some phases during the gait cycle that were clearly distinguishing between the different pathological patterns and TD gait, even though they were not specified by the classification rules. This was the case for the ankle kinematics during the first and third rocker. Specific characteristics of the first and third ankle rockers are currently not included in the pattern definitions of the Delphi study, nor were they commonly included in previously reported classifications[2,28,60–62]. During the reliability study (Chapter 4), several raters also indicated that the ankle joint waveforms of certain patients were unclassifiable. In these cases, raters felt that the first or third rocker was deviating to a degree they would not consider to be 'minor'. In addition, this issue had previously arisen during the Delphi study

(Chapter 3). After the third consensus round there was agreement that reduced ankle power generation during push-off (which can be expected to coexist with a decreased range of motion during the third ankle rocker) was an essential feature to include in the classification[8]. However, it was considered likely that ankle power generation could be observed in combination with several of the other ankle patterns that were already included, and there was no specific proposal on how this feature should be included in the final classification. Hence, further discussions and research are needed to assess the additional value of incorporating characteristics of the first and third ankle rocker into the classification system.

Seven patterns were defined for the knee during the stance phase and because of this high number of patterns, it was not surprising that the knee patterns during stance were found to have the highest amount of disagreement (Chapter 4). The level of clinician agreement on six out of seven patterns was found to be poor or moderate (kappa ranges between 0.32 and 0.46). Table 2 shows the confusion matrix of the knee during stance.

Table 2. Confusion matrix of the knee patterns during stance phase visualizes the confusion of the rater group (experienced (n=15) and inexperienced (n=14) raters) with respect to the criterion classification, based on the results from Chapter 4. For instance the cell in column KSTS2 and row KSTS1 indicates that 73 trials were classified as 'increased knee flexion at initial contact and earlier knee extension movement' (KSTS2) according to the criterion classification, while inexperienced or experienced clinical raters indicated they belonged to the pattern 'increased knee flexion at initial contact' (KSTS1). Cells in dark grey indicate most commonly observed disagreements; cells in light grey indicate likely accidental disagreements.

		Criterion classification						
		KSTS0	KSTS1	KSTS2	KSTS3	KSTS4	KSTS5	KSTS6
Rater group (n=29)	KSTS0	51	7	2	5	0	10	3
	KSTS1	11	63	73	0	7	24	14
	KSTS2	10	9	89	5	34	31	14
	KSTS3	13	1	0	54	9	0	0
	KSTS4	0	0	4	5	41	0	1
	KSTS5	12	2	5	0	5	85	14
	KSTS6	0	2	1	0	0	12	61

KSTS0 = no or minor gait deviations
KSTS1 = increased knee flexion at initial contact
KSTS2 = increased knee flexion at initial contact + earlier knee extension movement
KSTS3 = knee hyperextension
KSTS4 = knee hyperextension and increased knee flexion at initial contact
KSTS5 = increased knee flexion in midstance and internal flexion moment present
KSTS6 = increased knee flexion in midstance and internal extension moment present

The largest amount of confusion by far is caused by the pattern KSTS2, which is most often confused with KSTS1. A substantial amount of this confusion could be due to the continuous nature of the data. In addition, the rater group frequently assigned the pattern KSTS2 when the criterion classification identified a patient as KSTS4 or KSTS5. The confusion between these patterns suggests more strongly (as opposed to the confusion between KSTS1 and KSTS2) that pattern definitions for KSTS2, KSTS4, and KSTS5 might not be entirely clear. Regarding the knee patterns during stance, a concise learning phase might not suffice. It is advisable to provide several kinematic and kinetic examples of all knee patterns during the learning phase, as it was shown in Chapter 5 that on average, there are significantly large differences between the knee patterns for substantially wide parts of the stance phase (Cfr. Chapter 5, figure 5). On the other hand, the inclusion of KSTS1 and KSTS2 as separate patterns might be questioned for two reasons. Firstly, there was already a substantial amount of debate on the definition of the pattern KSTS2 during the Delphi study[8]. Secondly, Chapter 6 failed to show that the pattern KSTS2 was able to distinguish between clinically relevant subgroups in CP (Cfr. Chapter 6, table S3). Clinical experts should reconsider the added value of including KSTS2 in the classification.

The aforementioned reflections on the patterns of the hip in the coronal plane, as well as the knee patterns and ankle patterns during stance, show that much more can be learned from the results of Chapters 3-6 when they are combined, as opposed to when they are considered in isolation. However, it is beyond the scope of this general discussion to reflect on each pattern of the classification in detail. Feedback on these results should now be returned to the clinical experts, who should explore whether and how this information should be used to adapt the classification.

General methodological considerations

Readers should keep a few general methodological considerations in mind.

Firstly, the database that has been developed during the project was a retrospective sample of convenience from a hospital-based setting. Out of 2122 patients with CP that were present in the database of the clinical motion analysis laboratory of UZ Leuven in July 2016, 356 patients (17%), were selected who were between 3 and 18 years old, and who were diagnosed with spastic CP and GMFCS level I-III, either walking with or without walking aids. Out of all patients in the hospital database that were classified as ‘hemiplegic’ or ‘diplegic’, 19% and

22% respectively were recruited during the PhD research. On the other hand, only 4% and 2% of all patients classified as 'triplegia' and 'quadriplegia' were included. This might explain why patients with GMFCS level III were underrepresented in the recruited sample. Based on data from previous population-based studies, the distribution of GMFCS level I, II, and III is approximately 47%, 22%, and 31% respectively, whereas the distribution in the recruited database is 54% (level I), 33% (level II), and 13% (level III)[63,64]. It can be assumed that patients with GMFCS level III might be less often scheduled for an instrumented 3DGA, and are more often studied via observational video gait analysis. It also seems that patients with GMFCS level III were excluded more often because signs of experimental error (based on the range of motion and position of the knee varus-valgus angle) were more frequently detected. However, there is no objective data to support this claim.

Secondly, because patients are only recruited from a hospital setting, results are not generalizable towards the entire CP population. However, the hospital-based recruitment setting for these research studies was justified as instrumented biomechanical gait analyses are primarily performed to improve the understanding of gait pathology in CP and support treatment planning for patients who are being followed in a clinical setting. Hence, tools to interpret and analyze 3DGA data, such as gait classifications, are less relevant from an epidemiological point of view.

Thirdly, as was outlined in Chapter 6, surgical strategies have evolved rapidly during the past decades, which is a threat for the generalizability of retrospective research studies in general. A relatively large number of patients that were recruited underwent previous muscle tendon lengthening surgeries, an intervention which is currently performed much less frequently than approximately fifteen to twenty years ago[65,66]. However, quantitative data analysis techniques need large databases, which are difficult to obtain through the efforts of one research facility and therefore typically require retrospective data to be recruited. If larger multi-center databases are made publicly available, it will be easier to recruit a sufficiently large experimental patient population more restrictively (e.g. not include any patients with previous Achilles tendon lengthening surgery). An example of such a database is currently created by a European funded project (MD-Paedigree; http://www.md-paedigree.eu/). The aim of the project is to store, share, and analyze multi-center prospective and retrospective 3DGA data, varying from waveform data to mutually agreed-upon clinically relevant gait features. In addition to gait analysis data, variables from clinical examination as well as patient and

treatment history will also be incorporated. The clinical aim of this database is to perform 'similarity searches' to identify patients that are characterized by a gait pattern and a clinical background that is most similar to a patient being examined by the clinician. Studying treatment history and treatment outcomes of these similar patients may provide relevant information to support the clinical decision making for the individual new patient. These projects emphasize the need of anonymizing all patient data and implementing a data curation process in the development of research databases. A part of the curated and anonymized database composed during this PhD research has already been included in the MD Paedigree project and the full database might be uploaded in the future.

Considering the abovementioned issues regarding the patient selection for this doctoral dissertation, it can be hypothesized that the prevalence of the gait patterns (reported in Chapters 5-6), and the associations of the patterns with patient-specific variables (such as GMFCS, reported in Chapter 6), might slightly change if patient recruitment would have been performed prospectively and internationally across multiple research centers. However, associations with clinical characteristics (Chapter 6) and all other general conclusions for Chapters 2,4,5, and 6 are not likely to change given a wider, prospective patient recruitment.

Fourthly, inclusion of kinetic data in the classification system was only minimal, even though important pathological kinetic deviations and patterns in patients with CP have been described before[67]. However, it was found essential that the gait patterns were generalizable to all ambulatory children with CP. Since the use of walking aids in patients with GMFCS level III prohibits kinetic data analysis, patterns should not be solely dependent on kinetic features. In future research, a description of the kinetic behavior of the developed patterns could further fine-tune the classification. Such efforts have been reported before for the classification systems of Winters et al.[29] and Sutherland et al.[35 34,68].

Fifthly, walking velocity is a potential confounding factor that could influence a patient's classification. By interpreting time-normalized waveforms, this effect of walking velocity is slightly reduced, but cannot be ruled out. The effect of walking velocity on kinematic and kinetic data was not controlled for and it is unclear to what extent a patient's classification would change if patients walk faster or slower than their self-selected pace. It is not clear whether kinematic deviations arise due to a walking velocity that is different from TD children, or vice versa, because other factors influence the way in which joints move during gait, thereby causing walking velocity to change. It is relevant to evaluate patients at their

self-selected speed as this likely represents their typical behavior during daily activities. A patient's gait pattern arises due to a complex interplay of many variables and cannot be considered the direct representation of the brain lesions at the base of CP. Walking velocity, motivation, range of motion, bony deformities, fatigue, and possibly many other personal or environmental variables will all interact and contribute to specific deviations to different extents. The research within this PhD aimed to provide a tool to interpret 3DGA kinematics, and to a lesser extent kinetics, in a standardized and meaningful way so that it could be useful in clinical practice. It should be stressed that 3DGA is only a part of a patient's clinical examination and should thus be interpreted alongside other relevant parameters.

Lastly, the patterns of the classification were defined based on a threshold, which was the standard deviation of the average gait pattern of 56 TD children that were measured at the Clinical Motion Analysis Laboratory at University Hospitals Leuven. While the use of only one standard deviation as a threshold for pathological gait does not guarantee full exclusion of typical gait, it is representative of common clinical practice. However, differences between reference databases of TD children at different research centers can arise due to extrinsic measurement errors, as well as different measurement equipment or data processing techniques[69,70]. It is not yet clear to what extent results of the presented studies would change due to this issue.

Future research directions and conclusion

Future research directions

After the results of the presented research studies within this thesis have been discussed, and/or have led to appropriate adaptations to the gait patterns of the classification, several future research steps can be envisioned that will support the use of the patterns for research purposes as well as in clinical practice.

As outlined before, this doctoral research was part of a larger research project, funded by KU Leuven (OT/12/100). This larger research project aims to automate the developed classification using a probabilistic classification approach. A supervized, Bayesian approach can learn to classify kinematic and kinetic patient trials in the gait patterns as they were defined by the experts. A probabilistic approach tackles the abovementioned issues of intra-subject variability and expert disagreement due to the classification of continuous data,

by providing a probability score for belonging to each of the classes rather than outputting the most likely or probable class (Cfr. Psychometric properties, Reflections). In this way, a 'soft' and automated classification will be developed, providing clinicians with a pattern profile for each child. A Bayesian probabilistic approach has the additional advantage of being able to incorporate prior knowledge, which could be clinically inspired. The potential of probabilistic Bayesian approaches for classifying gait in CP has previously been shown by Van Gestel et al.[7], Di Lello et al.[71], and Zhang et al.[72]. The current PhD project established strong fundaments to explore a variety of supervized classification techniques to develop automatic classification algorithms by achieving an international consensus on clinically relevant gait patterns and by making an extended, classified, and well-structured database available. As part of the previously mentioned European MD Paedigree project, several probabilistic approaches for classifying CP gait are currently investigated.

Another crucial step forward that builds on the classification framework which has been developed in this PhD project entails the assessment of the responsiveness of the classification. Responsiveness or '*the ability of an instrument to detect change over time in the construct to be measured'* is a crucial psychometric property, which was not yet addressed in this thesis[73]. Prospective longitudinal intervention studies can now be organized to assess the responsiveness of the patterns to different treatment interventions, in relation to another comparator instrument, such as SPM or the Gait Profile Score. For such a study, careful consideration of what constitutes as success or improvement of gait pattern, is required. 'Success' in this case might not be equal to a change towards the patterns 'no or minor gait deviations'. Although the patterns of the classification represent a nominal scale, there are some patterns that show different levels of severity (similar to an ordinal ranking). For example at the level of the ankle during stance, 'equinus gait' could be considered a more severe form of 'reversed second ankle rocker', which could in turn be considered a more severe stage after 'horizontal second ankle rocker'. Patients changing from equinus gait to the pattern of a horizontal second ankle rocker post treatment might likely also be considered to have an improved gait pattern. In the undesirable event that the responsiveness of the patterns would be judged as 'poor', automatic probabilistic classifications will allow to estimate severity of gait pathology within each pattern.

Future research on the validity of the classification is also warranted. It is especially important that other research centers use the classification to consolidate both content and construct

validity, and extend the generalizability of the gait patterns. To do so, the online classification tool that was developed for the reliability study (www.cmal-tools-leuven.be) is ready to be used to facilitate the uptake of the classification by other clinical or research centers. Alongside the tool, guidelines for the classification can be uploaded along with training data. New raters can then use these data to learn and compare their results to expert classifications. The confusion matrices that are available from the reliability study provide further insight as to why experienced raters performed significantly better than inexperienced raters on some of the joints. This information can be used to further improve definitions or provide more elaborate pattern descriptions. If pattern definitions or elements of the classifications are adapted based on further research, these changes would then be reported and updated in guidelines. A similar approach has also been successfully applied to update the classification guidelines for the Gross Motor Function Classification System[39] and Manual Ability Classification System[74] based on findings from new research.

In the beginning of the project, it was decided to employ a comprehensive, 'bottom-up' approach, so the focus was directed at gait features and single joint patterns during (phases of) the gait cycle. This ensured that relevant deviations at each level would be taken into consideration, which may have facilitated the process of achieving a consensus. Future research can use the developed classification as a foundation to explore how joint patterns are combined into a total pattern at patient level (i.e. over different joints). Over 90% of all limbs that were included in the analysis of Chapter 6 were classified into one of the pathological patterns (i.e. not the patterns with 'no or minor gait deviations') for at least four out of eleven joints spread over the three anatomical planes. There was a median number of six 'pathological' joint patterns per limb. New study activities are currently exploring how these patterns at the level of different joints interact. As a first step, explorations are performed to assess 'simple multiple joint patterns', combining two joints or two phases (e.g. ankle and knee during stance phase). An interesting study in this respect was reported by Simon et al.[75], who investigated the distribution of kinematic transverse plane profiles in 188 children with spastic diplegia. As several lower limb muscles are bi-articular and therefore act around more than one joint, these explorations of simple multiple joint patterns may improve our understanding of the behavior of the underlying musculoskeletal model during gait. The search for 'total or global patterns' including all lower limb joints across multiple planes could also be performed. However, due to the high number of potential pattern combinations, this will likely cause several relevant joint patterns to be disregarded or a very high amount of

total patterns to be defined. It is expected that these total patterns will mainly have a descriptive role and will be less useful for clinicians in the clinical-decision-making process.

As was indicated in the consensus study (Chapter 3), the classification can be extended for instance with specific trunk or foot patterns. To do this, the workflow or methodological framework that was developed and applied in the current PhD project can be used as guidance. In this respect, a collaboration was set up with one of the former panel members of the consensus study, resulting in a Delphi consensus project that is currently ongoing in the UK to define foot patterns based on 3DGA data that are collected via multi-segment foot models.

Conclusion

The overall aim of this PhD research was to develop a clinically relevant, valid, and reliable classification system for pathological movement patterns during gait in children with CP, based on kinematic and kinetic data. Face validity (Chapter 3), content validity (Chapter 5), construct validity (Chapter 6), and reliability (Chapter 4) of the developed classification have been examined and demonstrated. The classification described within this PhD thesis presents clinicians and researchers with a comprehensive overview of clinically relevant features and joint patterns during gait, allowing a more uniform communication on gait pathology in CP. The patterns can be used in research to systematically describe, compare, and evaluate experimental patient populations, to investigate inter-joint relations during gait, or to further unravel the underlying mechanisms that cause pathological gait deviations to occur. In medical practice, clinicians can use the classification to track a patient's gait pattern over time or after treatment. In the future, the gait classification can be used alongside other standardized clinical measurements (e.g. electromyography data, instrumented weakness and spasticity measures), patient-specific characteristics, and brain imaging data, to establish the prognosis of treatment outcomes and directly support the interdisciplinary team in fine-tuning a patient's treatment.

In summary, this PhD research has taken important steps to improve the analysis and standardized interpretation of instrumented biomechanical analyses of gait in children with CP. Although the classification described within this thesis could already be used in practice, its responsiveness remains to be examined. It is important that the gait patterns and their definitions are adapted when necessary and that future studies from different clinical centers or research institutions further demonstrate the validity, responsiveness, and clinical

applicability of the classification. Specific web-based tools such as those that were developed during the PhD project can be applied to facilitate these future plans. In addition to the classification, SPM was identified as an objective, valid, and useful statistical approach to analyze kinematic and kinetic data, thereby providing several potential applications on its own. Among others, it can be used as an exploratory method to extract clinically relevant information from gait waveforms, especially for patient populations or medical conditions for which clinical expert knowledge is not readily available yet. It is important to consider that the applied methodological approach of this PhD thesis, combining qualitative and quantitative research methods to build a clinically relevant, reliable, and valid classification, could also be generalized to any other medical condition that affects movement.

References

1. Wren TAL, Rethlefsen S, Kay RM. Prevalence of specific gait abnormalities in children with cerebral palsy: influence of cerebral palsy subtype, age, and previous surgery. *J Pediatr Orthop*. 2005;25(1):79-83.
2. Dobson F, Morris ME, Baker R, Graham HK. Gait classification in children with cerebral palsy: A systematic review. *Gait Posture*. 2007;25(1):140-152.
3. Chambers H. The Delphi consensus technique: Oracle of gait analysis. *Dev Med Child Neurol*. 2016;58(3):228.
4. Skaggs DL, Rethlefsen S a, Kay RM, Dennis SW, Reynolds R a, Tolo VT. Variability in gait analysis interpretation. *J Pediatr Orthop*. 2000;20(6):759-764.
5. Chau T. A review of analytical techniques for gait data. Part 1: Fuzzy, statistical and fractal methods. *Gait Posture*. 2001;13(1):49-66.
6. Nieuwenhuys A, Papageorgiou E, Pataky T, De Laet T, Molenaers G, Desloovere K. Literature Review and Comparison of Two Statistical Methods to Evaluate the Effect of Botulinum Toxin Treatment on Gait in Children with Cerebral Palsy. *PLoS One*. 2016;11(3):e0152697.
7. Van Gestel L, De Laet T, Di Lello E, et al. Probabilistic gait classification in children with cerebral palsy: A Bayesian approach. *Res Dev Disabil*. 2011;32(6):2542-2552.
8. Nieuwenhuys A, Õunpuu S, Van Campenhout A, et al. Identification of joint patterns during gait in children with cerebral palsy: A Delphi consensus study. *Dev Med Child Neurol*. 2016;58(3):306-313.
9. Holm S. A Simple Sequentially Rejective Multiple Test Procedure. *Scand J Stat*. 1979;6(2):65-70.
10. Pataky TC. Generalized n-dimensional biomechanical field analysis using statistical parametric mapping. *J Biomech*. 2010;43(10):1976-1982.
11. Desloovere K, Molenaers G, Jonkers I, et al. A randomized study of combined botulinum toxin type A and casting in the ambulant child with cerebral palsy using objective outcome measures. *Eur J Neurol*. 2001;8 Suppl 5:75-87.
12. Molenaers G, Desloovere K, De Cat J, et al. Single event multilevel botulinum toxin type A treatment and surgery: similarities and differences. *Eur J Neurol*. 2001;8 Suppl 5:88-97.

13. Desloovere K, Molenaers G, De Cat J, et al. Motor function following multilevel botulinum toxin type A treatment in children with cerebral palsy. *Dev Med Child Neurol*. 2007;49(1):56-61.

14. Metaxiotis D, Siebel A, Doederlein L. Repeated botulinum toxin A injections in the treatment of spastic equinus foot. *Clin Orthop Relat Res*. 2002;(394):177-185.

15. Polak F, Morton R, Ward C, Wallace WA, Doderlein L, Siebel A. Double-blind comparison study of two doses of botulinum toxin A injected into calf muscles in children with hemiplegic cerebral palsy. *Dev Med Child Neurol*. 2002;44(8):551-555.

16. Pataky TC, Vanrenterghem J, Robinson MA. The probability of false positives in zero-dimensional analyses of one-dimensional kinematic, force and EMG trajectories. *J Biomech*. 2016;49:1468-1476.

17. Baker R, McGinley JL, Schwartz MH, et al. The gait profile score and movement analysis profile. *Gait Posture*. 2009;30(3):265-269.

18. Adler RJ, Taylor JE. *Random Fields and Geometry*. Vol 1st ed. New York: Springer-Verlag New York; 2007.

19. Pataky TC, Vanrenterghem J, Robinson MA. Zero- vs. one-dimensional, parametric vs. non-parametric, and confidence interval vs. hypothesis testing procedures in one-dimensional biomechanical trajectory analysis. *J biomech*. 2015;48(ii):1277-1285.

20. Pataky TC, Robinson M a., Vanrenterghem J. Vector field statistical analysis of kinematic and force trajectories. *J Biomech*. 2013;46(14):2394-2401.

21. Knudson D, Lindsey C. Type I and type II errors in correlations of various sample sizes. *Compr Psychol*. 2014;3(1):811-814.

22. Oeffinger D, Bagley A, Rogers S, et al. Outcome tools used for ambulatory children with cerebral palsy. *Dev Med Child Neurol*. 2008;50:918-925.

23. Klejman S, Andrysek J, Dupuis A, Wright V. Test-retest reliability of discrete gait parameters in children with cerebral palsy. *Arch Phys Med Rehabil*. 2010;91(5):781-787.

24. Sutherland DH, Kaufman KR, Wyatt MP, Chambers HG, Mubarak SJ. Double-blind study of botulinum A toxin injections into the gastrocnemius muscle in patients with cerebral palsy. *Gait Posture*. 1999;10(1):1-9.

25. McGinley JL, Baker R, Wolfe R, Morris ME. The reliability of three-dimensional kinematic gait measurements: A systematic review. *Gait Posture*. 2009;29(3):360-369.

26. Kaufman K, Miller E, Kingsbury T, et al. Reliability of 3D Gait Data across Multiple Laboratories. *Gait Posture*. 2016;49:375-381.

27. Hasson F, Keeney S, McKenna H. Research guidelines for the Delphi survey technique. *J Adv Nurs*. 2000;32(4):1008-1015.

28. Davids JR, Bagley AM. Identification of Common Gait Disruption Patterns in Children With Cerebral Palsy. *J Am Acad Orthop Surg*. 2014;22(12):782-790.

29. Winters TF, Gage JR, Hicks R. Gait patterns in spastic hemiplegia in children and young adults. *J Bone Joint Surg Am*. 1987;69:437-441.

30. Ferrari A, Alboresi S, Muzzini S, Pascale R, Perazza S, Cioni G. The term diplegia should be enhanced. Part I: a new rehabilitation oriented classification of cerebral palsy. *Eur J Phys Rehabil Med*. 2008;44(2):195-201.

31. Rodda JM, Graham HK, Carson L, Galea MP, Wolfe R. Sagittal gait patterns in spastic diplegia. *J Bone Joint Surg Br*. 2004;86(2):251-258.

32. Simon S, Deutsch S, Nuzzo R, et al. Genu recurvatum in spastic cerebral palsy. *J Bone Joint*

Surg Am. 1978;60-A:882-894.

33. Hullin MG, Robb JE, Loudon IR. Gait patterns in children with hemiplegic spastic cerebral palsy. *J Pediatr Orthop B*. 1996;5(4):247-251.

34. Lin C, Guo L, Su F, Chou Y, Cherng R. Common abnormal kinetic patterns of the knee in gait in spastic diplegia of cerebral palsy. *Gait Posture*. 2000;11:224-232.

35. Sutherland DH, Davids JR. Common Gait Abnormalities of the Knee in Cerebral Palsy. *Clin Orthop Relat Res*. 1993;288:139-147.

36. Remijn L, Speyer R, Groen BE, Holtus PCM, van Limbeek J, Nijhuis-van der Sanden MWG. Assessment of mastication in healthy children and children with cerebral palsy: A validity and consistency study. *J Oral Rehabil*. 2013;40(5):336-347.

37. Sellers D, Mandy A, Pennington L, Hankins M, Morris C. Development and reliability of a system to classify the eating and drinking ability of people with cerebral palsy. *Dev Med Child Neurol*. 2014;56(3):245-251.

38. Hidecker M, Paneth N, Rosenbaum P, et al. Developing and validating the communication function classification system for individuals with cerebral palsy. *Dev Med Child Neurol*. 2012;53(8):704-710.

39. Palisano R, Rosenbaum P, Walter S, Russell D, Wood E, Galuppi B. Development and reliability of a system to classify gross motor function in children with cerebral palsy. *Dev Med Child Neurol*. 1997;39(2):214-223.

40. Morris C, Bartlett DJ. Gross Motor Function Classification System : impact and utility. *Dev Med Child Neurol*. 2004;46:60-65.

41. Agresti A. Inference for contingency tables. In: *Categorical Data Analysis*. Vol ; 2002:70-114.

42. Jones J, Hunter D. Qualitative Research: Consensus methods for medical and health services research. *BMJ*. 1995;311:376-380.

43. Delbecq A, Van De Ven A. A group process model for problem identification and program planning. *J Appl Behav Sci*. 1971;7(4):466-492.

44. McMillan SS, King M, Tully MP. How to use the nominal group and Delphi techniques. *Int J Clin Pharm*. 2016;38(3):1-8.

45. Keeney S, Hasson F, McKenna H. Consulting the oracle: Ten lessons from using the Delphi technique in nursing research. *J Adv Nurs*. 2006;53(2):205-212.

46. Dobson F, Morris M, Baker R, Wolfe R, Graham H. Clinician agreement on gait pattern ratings in children with spastic hemiplegia. *Dev Med Child Neurol*. 2006;48(6):429-435.

47. Stott NS, Atherton WG, Mackey AH, Galley IJ, Nicol RO, Walsh SJ. The reliability and validity of assessment of sagittal plane deviations in children who have spastic diplegia. *Arch Phys Med Rehabil*. 2005;86(12):2337-2341.

48. McDowell BC, Kerr C, Kelly C, Salazar J, Cosgrove A. The validity of an existing gait classification system when applied to a representative population of children with hemiplegia. *Gait Posture*. 2008;28(3):442-447.

49. Riad J, Haglund-Akerlind Y, Miller F. Classification of Spastic Hemiplegic Cerebral Palsy in Children. *J Pediatr Orthop*. 2007;27(7):758-764.

50. de Morais Filho MC, Kawamura CMC, Lopes JJAF, Neves DLD, Cardoso MM de O, Caiafa JJB. Most frequent gait patterns in diplegic spastic cerebral palsy. *Acta Orthop Bras [online]*. 2014;22(4):197-201.

51. Dobson F, Morris ME, Baker R, Graham HK. Unilateral cerebral palsy: a population-based study of gait and motor function. *Dev Med Child Neurol*. 2011;53(5):429-435.

52. Cioni G, Lodesani M, Pascale R, et al. The term diplegia should be enhanced. Part II: Contribution to validation of the new rehabilitation oriented classification. *Eur J Phys Rehabil Med*. 2008;44(2):203-211.

53. Rozumalski A, Schwartz MH. Crouch gait patterns defined using k-means cluster analysis are related to underlying clinical pathology. *Gait Posture*. 2009;30(2):155-160.

54. Schless SH, Desloovere K, Aertbeliën E, Molenaers G, Huenaerts C, Bar-On L. The intra- and inter-rater reliability of an instrumented spasticity assessment in children with cerebral palsy. *PLoS One*. 2015;10(7):1-23.

55. Schwartz MH, Trost JP, Wervey R a. Measurement and management of errors in quantitative gait data. *Gait Posture*. 2004;20(2):196-203.

56. Steinwender G, Saraph V, Scheiber S, Zwick EB, Uitz C, Hackl K. Intrasubject repeatability of gait analysis data in normal and spastic children. *Clin Biomech*. 2000;15(2):134-139.

57. Redekop S, Andrysek J, Wright V. Single-session reliability of discrete gait parameters in ambulatory children with cerebral palsy based on GMFCS level. *Gait Posture*. 2008;28(4):627-633.

58. Baker R. Why bother with gait classification? https://wwrichard.net/2013/06/06/why-bother-with-gait-classification/.

59. Tirosh O, Dobson F, Graham K, Baker R. Gait classification - a step in the wrong direction. *Dev Med Child Neurol*. 2009;51(suppl. 2):1-90.

60. Carriero A, Zavatsky A, Stebbins J, Theologis T, Shefelbine SJ. Determination of gait patterns in children with spastic diplegic cerebral palsy using principal components. *Gait Posture*. 2009;29(1):71-75.

61. Toro B, Nester CJ, Farren PC. Cluster analysis for the extraction of sagittal gait patterns in children with cerebral palsy. *Gait Posture*. 2007;25(2):157-165.

62. Bonnefoy-Mazure A, Sagawa YJ, Lascombes P, De Coulon G, Armand S. Identification of gait patterns in individuals with cerebral palsy using multiple correspondence analysis. *Res Dev Disabil*. 2013;34(9):2684-2693.

63. Gorter JW, Rosenbaum PL, Hanna SE, et al. Limb distribution, motor impairment, and functional classification of cerebral palsy. *Dev Med Child Neurol*. 2004;46:461-467.

64. Rosenbaum PL. Prognosis for Gross Motor Function in Cerebral Palsy: Creation of Motor Development Curves. *JAMA J Am Med Assoc*. 2002;288(11):1357-1363.

65. Laracca E, Stewart C, Postans N, Roberts A. The effects of surgical lengthening of hamstring muscles in children with cerebral palsy - The consequences of pre-operative muscle length measurement. *Gait Posture*. 2014;39(3):847-851.

66. Gage JR. Con:Interobserver variability of gait analysis. *J Pediatr Orthop*. 2003;23(6):290-291.

67. Gage J, Schwartz M, Koop S, Novacheck T. *DThe Identification and Treatment of Gait Problems in Cerebral Palsy*. Vol 2nd editio. London: Mac Keith Press; 2009.

68. Stout J, Bruce B, Gage J, Schutte L. Joint kinetic patterns in children with spastic hemiplegia cerebral palsy. *Gait Posture*. 1995;3(4):274.

69. Pinzone O, Schwartz M, Thomason P, Baker R. The comparison of normative reference data from different gait analysis services. *Gait Posture*. 2014;40(2):286-290.

70. Chester VL, Tingley M, Biden EN. Comparison of two normative paediatric gait databases. *Dyn Med*. 2007;6:8.

71. Di Lello E. Bayesian Time-Series Models : Expert Knowledge-Driven Inference and Learning for Engineering Applications. 2015;(February).

72. Zhang B, Zhang Y, Begg RK. Gait classification in children with cerebral palsy by Bayesian approach. *PATTERN Recognit*. 2009;42(4, SI):581-586.

73. Mokkink LB, Terwee CB, Patrick DL, et al. The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes. *J Clin Epidemiol*. 2010;63(7):737-745.

74. Eliasson A-C, Krumlinde-Sundholm L, Rösblad B, et al. The Manual Ability Classification System (MACS) for children with cerebral palsy: scale development and evidence of validity and reliability. *Dev Med Child Neurol*. 2006;48(7):549-554.

75. Simon A-L, Ilharreborde B, Megrot F, et al. A Descriptive Study of Lower Limb Torsional Kinematic Profiles in Children With Spastic Diplegia. *J Pediatr Orthop*. 2015;35(6):576-582.

Appendices

Appositions

About the author

Bibliography

Author affiliations

Appositions

Het belang van een muzikale opleiding voor de persoonlijke ontwikkeling van elk individu moet vaker worden benadrukt.

Een inschatting van de economische impact van pestgedrag op de werkvloer zal een stimulans vormen om de taboesfeer rond deze multifactoriële problematiek verder op te heffen en vooruitgang te boeken in het streven naar een pestvrije werkomgeving.

Het democratisch bestuursmodel wordt weleens beschouwd als een slechte vorm van bestuur, maar beter dan alle anderen. Dit geldt ook voor het systeem van enkel- of dubbelblinde peer-review, dat door wetenschappelijke tijdschriften gehanteerd wordt als garantie voor kwalitatief, hoogstaand onderzoek.

About the author

Angela Nieuwenhuys was born on October 10, 1989 and grew up in Booischot, a small village about 30km from Leuven, Belgium. After finishing secondary school, she had to make a tough choice between a professional career in music or science. Her fascination with the human body, its anatomy and the way it functions, drove her to start the Bachelor program in Physical Therapy at KU Leuven in 2007. She combined her studies with a position as an oboist in the University Symphony Orchestra of KU Leuven, under the professional guidance of Edmondus Saveniers. She graduated Magna cum Laude in 2012, obtaining a Master of Science degree in Rehabilitation Sciences and Physiotherapy. During her classes and internships, she became aware that the evidence-base for many physiotherapeutic interventions remains small. This inspired and drove her towards academic research. Three months after her graduation, she was given the opportunity to start a PhD project under the supervision of Prof. Kaat Desloovere, Prof. Ir. Tinne De Laet, and Prof. Dr. Guy Molenaers. This PhD thesis was part of a larger project, funded by KU Leuven and entitled: 'Pattern recognition for gait analysis: integration of clinical expert knowledge and machine learning techniques' (OT/12/100). Apart from her professional endeavors, Angela is a passionate, competitive badminton player. She also enjoys concerts and a nice glass of wine with friends.

Bibliography

IT (Articles in internationally reviewed academic journals)

Nieuwenhuys, A., Papageorgiou E., Molenaers, G., De Laet, T., Desloovere, K. Prevalence of joint patterns during gait in children with cerebral palsy is related to gross motor function level, topographical classification, weakness, and spasticity: an exploratory study. Frontiers in Human Neuroscience [in review]

Nieuwenhuys, A., Papageorgiou E., Desloovere, K., Molenaers, G., De Laet, T. Statistical parametric mapping to identify differences between consensus-based gait patterns in children with cerebral palsy. [in preparation for Journal of Biomechanics]

Santos LG, **Nieuwenhuys, A.,** Monari, D., Desloovere, K, Russo LT. Features and statistical parametric mapping analysis for kinematic data during drinking task of chronic hemiparetic subjects. [in preparation]

Meyer, C., Wesseling, M., Corten, K., **Nieuwenhuys, A.,** Monari, D., Simon J-P., Jonkers, K., Desloovere, K. Hip movement pathomechanics of patients with hip osteoarthritis aim at reducing hip joint loading at the osteoarthritic side. [in preparation for Arthritis Care and Research]

Nieuwenhuys, A., Papageorgiou, E., Molenaers, G., Monari, D., De Laet, T., Desloovere, K. (2016). Inter- and intrarater clinician agreement on joint patterns during gait in children with cerebral palsy. Developmental Medicine and Child Neurology. [in revision]

Nieuwenhuys, A., Papageorgiou, E., Pataky, T., De Laet, T., Molenaers, G., Desloovere, K. (2016). Literature review and comparison of two statistical methods to evaluate the effect of Botulinum Toxin treatment on gait in children with cerebral palsy. *PLoS One*. March 31; 10.1371/journal. pone.0152697.

Nieuwenhuys, A., Õunpuu, S., Van Campenhout, A., Theologis, T., De Cat, J., Stout, J., Molenaers, G., De Laet, T., Desloovere, K. (2016). Identification of joint patterns during gait in children with cerebral palsy: a Delphi consensus study. Developmental Medicine and Child Neurology, 58(3), 306-313.

Bar-On, L., Van Campenhout, A., Desloovere, K., Aertbeliën, E., Huenaerts, C., Vandendoorent, B., **Nieuwenhuys, A.,** Molenaers, G. (2014). Is an instrumented spasticity assessment an improvement over clinical spasticity scales in assessing and predicting the response to integrated Botulinum Toxin-A treatment in children with Cerebral Palsy?. Archives of Physical Medicine and Rehabilitation, 95 (3), 515-523.

Bar-On, L., Aertbeliën, E., Molenaers, G., Van Campenhout, A., Vandendoorent, B., **Nieuwenhuys, A.,** Jaspers, E., Hunaerts, C., Desloovere, K. (2014). Instrumented assessment of the effect of Botulinum Toxin-A in the medial hamstrings in children with cerebral palsy. Gait & posture, 39 (1), 17-22.

Heyrman, L., Feys, H., Molenaers, G., Jaspers, E., Monari, D., **Nieuwenhuys, A.,** Desloovere, K. (2014). Altered trunk movements during gait in children with spastic diplegia: compensatory or underlying trunk control deficit?. Research in Developmental Disabilities, 35 (9), 2044-2052.

Baert, I., Mahmoudian, A., **Nieuwenhuys, A.,** Jonkers, I., Staes, F., Luyten, F., Truijen, S., Verschueren, S. (2013). Proprioceptive accuracy in women with early and established knee osteoarthritis and its relation to functional ability, postural control, and muscle strength. Clinical Rheumatology, 32 (9), 1365-1374.

IMa (Meeting abstracts, presented at international scientific conferences and symposia, published or not published in proceedings or journals)

Nieuwenhuys, A., Papageorgiou, E., Molenaers, G., De Laet, T., Desloovere, K. (2016) Prevalence of joint patterns during gait in children with cerebral palsy is related to gross motor function level, topographical classification, strength and spasticity. *Gait & Posture*. Annual meeting of the European Movement Analysis Society for Adults and Children. Sevilla, Spain, 28 September - 1 October 2016.

Papageorgiou E., **Nieuwenhuys, A.,** Goudriaan, G., Van Campenhout, A., Ortibus, E., Molenaers, G., Desloovere, K. (2016) Gait pattern changes in children with cerebral palsy after the administration of botulinum toxin injections. *Gait & Posture*. Annual meeting of the European Movement Analysis Society for Adults and Children. Sevilla, Spain, 28 September - 1 October, 2016.

Van de Walle, P., Goudriaan M., **Nieuwenhuys, A.,** Callewaert, B., Desloovere, K., Hallemans, A. (2016) Age-related changes in continuous kinematic and kinetic gait curves in typical developing children and adults. *Gait & Posture*. Annual meeting of the European Movement Analysis Society for Adults and Children. Sevilla, Spain, 28 September - 1 October, 2016.

Nieuwenhuys, A., Papageorgiou, E., Dimitriou, G., Molenaers, G., De Laet, T., Desloovere, K. (2015). Inter- and intrarater reliability of an international consensus-based gait classification system in children with cerebral palsy. *Gait & Posture: Vol. 42S*. Annual meeting of the European Movement Analysis Society for Adults and Children. Heidelberg, Germany, 10-12 September 2015 (pp. S46).

Huenaerts, C., Molenaers, G., **Nieuwenhuys, A.,** Pauwels, P., Monari, D., Desloovere, K.. (2015). Selective dorsal rhizotomy with or without correction of bony deformities in children with cerebral palsy: the effect on gait. *Gait & Posture*: vol 42S. Annual meeting of the European Movement Analysis Society for Adults and Children. Heidelberg, Germany, 10-12 September 2015 (pp. S1-S101).

Nieuwenhuys, A., Monari, D., Deschamps, K., De Laet, T., Molenaers, G., Desloovere, K. (2014). Statistical parametric mapping to identify gait pattern features in children with cerebral palsy. *Gait & Posture.* Annual meeting of the European Movement Analysis Society for Adults and Children. Rome, Italy, 29 September - 4 October 2014.

Goudriaan, M., Bruyninckx, C., **Nieuwenhuys, A.,** Delanghe, K., Casie, J., Monari, D., Desloovere, K. (2014). *Gait & Posture.* Annual meeting of the European Movement Analysis Society for Adults and Children. Rome, Italy, 29 September - 4 October 2014.

Di Lello, E., **Nieuwenhuys, A.,** Desloovere, K., De Laet, T. (2014). Bayesian Hidden Markov Models for Segmentation of Gait Motion Capture Data. 1st Clinical Movement Analysis World Conference (ESMAC-SIAMOC). Rome, 29 September - 4 October 2014.

Di Lello, E., **Nieuwenhuys, A.,** De Laet, T., Desloovere, K. (2014). Identification of Gait Events Combining Bayesian Hidden Markov Models and Linear Regression. Workshop on latest advances on Natural Motion Understanding and Human Motion Synthesis. IEEE International Conference on Robotics and Automation. Hong Kong, 31 May - 7 June 2014.

Nieuwenhuys, A., Meyer, C., Monari, D. De Laet, T., Van Gestel, L., Molenaers, G., Desloovere, K. (2013). An innovative approach toward gait feature detection in children with CP. *Gait & Posture: Vol. 39S*. Annual meeting of the European Movement Analysis Society for Adults and Children. Glasgow, Scotland, 2-7 September 2013 (pp. S86-S87).

Desloovere, K., **Nieuwenhuys, A.,** Wambacq, H., Monari, D., Bar-On, L., Molenaers, G. (2013). Integrated instrumented weakness assessment to characterise lower limb strength in children with CP. *Gait & Posture: Vol. 39S*. Annual meeting of the European Movement Analysis Society for Adults and Children. Glasgow, Scotland, 2-7 September 2013 (pp. S1-S141). Gait & Posture 39S (2014) S1–S141: Butterworth-Heinemann.

Wambacq, H., Aertbeliën, E., Bar-On, L., Desloovere, K., **Nieuwenhuys, A.,** Molenaers, G. (2013). Advanced EMG analysis during gait: the probability of onset deviation index. *Gait & Posture: Vol. 39S*. Annual meeting of the European Movement Analysis Society for Adults and Children. Glasgow, Scotland, 2-7 September 2013 (pp. S1-S141). Gait & Posture 39S (2014) S1–S141: Butterworth-Heinemann.

Author affiliations

Corresponding author

Angela Nieuwenhuys

KU Leuven, Department of Rehabilitation Sciences, Leuven, Belgium.

angelanieuwenhuys@gmail.com

Co-authors

Kaat Desloovere

KU Leuven, Department of Rehabilitation Sciences, Leuven, Belgium.

Clinical Motion Analysis Laboratory, University Hospitals Leuven, Leuven, Belgium.

kaat.desloovere@uzleuven.be

Josse De Cat

KU Leuven, Department of Rehabilitation Sciences, Leuven, Belgium.

Clinical Motion Analysis Laboratory, University Hospitals Leuven, Leuven, Belgium.

josse.decat@kuleuven.be

Tinne De Laet

KU Leuven, Faculty of Engineering Science, Leuven, Belgium.

tinne.delaet@kuleuven.be

Guy Molenaers

KU Leuven, Department of Development and Regeneration, Leuven, Belgium.

Department of Orthopedics, University Hospitals Leuven, Leuven, Belgium.

guy.molenaers@uzleuven.be

Davide Monari

KU Leuven, Department of Mechanical Engineering, Division PMA, Leuven, Belgium.

davide.monari@kuleuven.be

Sylvia Õunpuu

Center for Motion Analysis, Division of Orthopaedics, Connecticut Children's Medical Center, Farmington, USA.

sounpuu@ccmckids.org

Eirini Papageorgiou

KU Leuven, Department of Rehabilitation Sciences, Leuven, Belgium.

eirini.papageorgiou@kuleuven.be

Todd Pataky

Department of Bioengineering, Shinshu University, Ueda, Japan.

tpataky@shinshu-u.ac.jp

Tim Theologis

Nuffield Orthopaedic Centre, Oxford University Hospitals NHS Trust, Oxford, UK.

tim.theologis@ouh.nhs.uk

Jean Stout

Gillette Children's Specialty Healthcare, St. Paul, USA.

jstout@gillettechildrens.com

Anja Van Campenhout

KU Leuven, Department of Development and Regeneration, Leuven, Belgium.

Department of Orthopedics, University Hospitals Leuven, Leuven, Belgium.

anja.vancampenhout@uzleuven.be